The Circuits and Filters Handbook

Third Edition

Analog and VLSI Circuits

Edited by

Wai-Kai Chen

CRC Press
Taylor & Francis Group

VISIT…

Analog and VLSI Circuits

The Circuits and Filters Handbook

Third Edition

Edited by

Wai-Kai Chen

Fundamentals of Circuits and Filters

Feedback, Nonlinear, and Distributed Circuits

Analog and VLSI Circuits

Computer Aided Design and Design Automation

Passive, Active, and Digital Filters

The Circuits and Filters Handbook

Third Edition

Analog and VLSI Circuits

Edited by

Wai-Kai Chen

University of Illinois

Chicago, U. S. A.

CRC Press

Taylor & Francis Group

Boca Raton London New York

CRC Press is an imprint of the

Taylor & Francis Group, an **informa** business

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

© 2009 by Taylor & Francis Group, LLC
CRC Press is an imprint of Taylor & Francis Group, an Informa business

No claim to original U.S. Government works
Printed in the United States of America on acid-free paper
10 9 8 7 6 5 4 3 2 1

International Standard Book Number-13: 978-1-4200-5891-8 (Hardcover)

This book contains information obtained from authentic and highly regarded sources. Reasonable efforts have been made to publish reliable data and information, but the author and publisher cannot assume responsibility for the validity of all materials or the consequences of their use. The authors and publishers have attempted to trace the copyright holders of all material reproduced in this publication and apologize to copyright holders if permission to publish in this form has not been obtained. If any copyright material has not been acknowledged please write and let us know so we may rectify in any future reprint.

Except as permitted under U.S. Copyright Law, no part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access www.copyright.com (http://www.copyright.com/) or contact the Copyright Clearance Center, Inc. (CCC), 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Library of Congress Cataloging-in-Publication Data

Analog and VLSI circuits / edited by Wai-Kai Chen.
p. cm.
Includes bibliographical references and index.
ISBN-13: 978-1-4200-5891-8
ISBN-10: 1-4200-5891-6
1. Linear integrated circuits. 2. Integrated circuits--Very large scale integration. 3. Electronic circuits. I. Chen, Wai-Kai, 1936- II. Title.

TK7874.654.A47 2009
621.39'5--dc22 2008048128

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Contents

Preface vii

Editor-in-Chief ix

Contributors xi

SECTION I Analog Integrated Circuits

1 Monolithic Device Models **1**-1
Bogdan M. Wilamowski, Guofu Niu, John Choma, Jr., Stephen I. Long, Nhat M. Nguyen, and Martin A. Brooke

2 Analog Circuit Cells **2**-1
Kenneth V. Noren, John Choma, Jr., J. Trujillo, David G. Haigh Bill Redman-White, Rahim Akbari-Dilmaghani, Mohammed Ismail, Shu-Chuan Huang, Chung-Chih Hung, and Trond Saether

3 High-Performance Analog Circuits **3**-1
Chris Toumazou, Alison Payne, John Lidgey, Alicja Konczakowska, and Bogdan M. Wilamowski

4 RF Communication Circuits **4**-1
Michiel Steyaert, Wouter De Cock, and Patrick Reynaert

5 PLL Circuits **5**-1
Muh-Tian Shiue and Chorng-Kuang Wang

6 Synthesis of Reactance Pulse-Forming Networks **6**-1
Igor M. Filanovsky

SECTION II The VLSI Circuits

7 Fundamentals of Digital Signal Processing **7**-1
Roland Priemer

8 Digital Circuits **8**-1
John P. Uyemura, Robert C. Chang, and Bing J. Sheu

9 Digital Systems **9**-1
Festus Gail Gray, Wayne D. Grover, Josephine C. Chang, Bing J. Sheu Roland Priemer, Kung Yao, and Flavio Lorenzelli

10 Data Converters **10**-1
Bang-Sup Song and Ramesh Harjani

Index **IN**-1

Preface

The purpose of this book is to provide in a single volume a comprehensive reference work covering the broad spectrum of monolithic device models, high-performance analog circuits, radio-frequency communications and PLL circuits, digital systems, and data converters. This book is written and developed for the practicing electrical engineers and computer scientists in industry, government, and academia. The goal is to provide the most up-to-date information in the field.

Over the years, the fundamentals of the field have evolved to include a wide range of topics and a broad range of practice. To encompass such a wide range of knowledge, this book focuses on the key concepts, models, and equations that enable the design engineer to analyze, design, and predict the behavior of large-scale circuits and systems. While design formulas and tables are listed, emphasis is placed on the key concepts and theories underlying the processes.

This book stresses fundamental theories behind professional applications and uses several examples to reinforce this point. Extensive development of theory and details of proofs have been omitted. The reader is assumed to have a certain degree of sophistication and experience. However, brief reviews of theories, principles, and mathematics of some subject areas are given. These reviews have been done concisely with perception.

The compilation of this book would not have been possible without the dedication and efforts of Professor John Choma, Jr., and most of all the contributing authors. I wish to thank them all.

Wai-Kai Chen

Editor-in-Chief

Wai-Kai Chen is a professor and head emeritus of the Department of Electrical Engineering and Computer Science at the University of Illinois at Chicago. He received his BS and MS in electrical engineering at Ohio University, where he was later recognized as a distinguished professor. He earned his PhD in electrical engineering at the University of Illinois at Urbana–Champaign.

Professor Chen has extensive experience in education and industry and is very active professionally in the fields of circuits and systems. He has served as a visiting professor at Purdue University, the University of Hawaii at Manoa, and Chuo University in Tokyo, Japan. He was the editor-in-chief of the *IEEE Transactions on Circuits and Systems*, *Series I and II*, the president of the IEEE Circuits and Systems Society, and is the founding editor and the editor-in-chief of the *Journal of Circuits, Systems and Computers*. He received the Lester R. Ford Award from the Mathematical Association of America; the Alexander von Humboldt Award from Germany; the JSPS Fellowship Award from the Japan Society for the Promotion of Science; the National Taipei University of Science and Technology Distinguished Alumnus Award; the Ohio University Alumni Medal of Merit for Distinguished Achievement in Engineering Education; the Senior University Scholar Award and the 2000 Faculty Research Award from the University of Illinois at Chicago; and the Distinguished Alumnus Award from the University of Illinois at Urbana–Champaign. He is the recipient of the Golden Jubilee Medal, the Education Award, and the Meritorious Service Award from the IEEE Circuits and Systems Society, and the Third Millennium Medal from the IEEE. He has also received more than a dozen honorary professorship awards from major institutions in Taiwan and China.

A fellow of the Institute of Electrical and Electronics Engineers (IEEE) and the American Association for the Advancement of Science (AAAS), Professor Chen is widely known in the profession for the following works: *Applied Graph Theory* (North-Holland), *Theory and Design of Broadband Matching Networks* (Pergamon Press), *Active Network and Feedback Amplifier Theory* (McGraw-Hill), *Linear Networks and Systems* (Brooks/Cole), *Passive and Active Filters: Theory and Implements* (John Wiley), *Theory of Nets: Flows in Networks* (Wiley-Interscience), *The Electrical Engineering Handbook* (Academic Press), and *The VLSI Handbook* (CRC Press).

Contributors

Rahim Akbari-Dilmaghani
Department of Electronic and Electrical Engineering
University College of London
London, United Kingdom

Martin A. Brooke
School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, Georgia

Josephine C. Chang
Ming Hsieh Department of Electrical Engineering
University of Southern California
Los Angeles, California

Robert C. Chang
Ming Hsieh Department of Electrical Engineering
University of Southern California
Los Angeles, California

John Choma, Jr.
Ming Hsieh Department of Electrical Engineering
University of Southern California
Los Angeles, California

Wouter De Cock
Department of Electrical Engineering
Catholic University of Leuven
Leuven, Belgium

Igor M. Filanovsky
Department of Electrical Engineering
University of Alberta
Edmonton, Alberta, Canada

Festus Gail Gray
Department of Electrical and Computer Engineering
Virginia Polytechnic Institute and State University
Blacksburg, Virginia

Wayne D. Grover
Network Systems
TRLabs
Edmonton, Alberta, Canada

and

Department of Electrical and Computer Engineering
University of Alberta
Edmonton, Alberta, Canada

David G. Haigh
Department of Electronic and Electrical Engineering
University College of London
London, United Kingdom

Ramesh Harjani
Department of Electrical Engineering
University of Minnesota
Minneapolis, Minnesota

Shu-Chuan Huang
Department of Electrical Engineering
Ohio State University
Columbus, Ohio

Chung-Chih Hung
Department of Electrical Engineering
Tatung Institute of Technology
Taipei, Taiwan

Mohammed Ismail
Department of Electrical Engineering
Ohio State University
Columbus, Ohio

Alicja Konczakowska
Department of Optoelectronics and Electronics Systems
Gdansk University of Technology
Gdansk, Poland

John Lidgey
School of Technology
Oxford Brookes University
London, United Kingdom

Stephen I. Long
Department of Electrical and Computer Engineering
University of California, Santa Barbara
Santa Barbara, California

Flavio Lorenzelli
SGS-Thomson Microelectronics
Milan, Italy

and

University of Milan, Crema
Crema, Italy

Nhat M. Nguyen
Rambus Inc.
Los Altos, California

Guofu Niu
Department of Electrical and Computer Engineering
Auburn University
Auburn, Alabama

Kenneth V. Noren
Department of Electrical and Computer Engineering
University of Idaho
Moscow, Idaho

Alison Payne
Institute of Biomedical Engineering
Imperial College of Science, Technology, and Medicine
London, United Kingdom

Roland Priemer
Department of Electrical and Computer Engineering
University of Illinois at Chicago
Chicago, Illinois

Bill Redman-White
School of Electronics and Computer Science
University of Southampton
Southampton, United Kingdom

Patrick Reynaert
Department of Electrical Engineering
Catholic University of Leuven
Leuven, Belgium

Trond Saether
Nordic VLSI A/S
Flatasen, Norway

Bing J. Sheu
Taiwan Semiconductor Manufacturing Company
Hsin-Chu, Taiwan

Muh-Tian Shiue
Department of Electrical Engineering
National Central University
Chung-Li, Taiwan

Bang-Sup Song
Department of Electrical and Computer Engineering
University of California, San Diego
San Diego, California

Michiel Steyaert
Department of Electrical Engineering
Catholic University of Leuven
Leuven, Belgium

Chris Toumazou
Institute of Biomedical Engineering
Imperial College of Science, Technology, and Medicine
London, United Kingdom

J. Trujillo
Ming Hsieh Department of Electrical Engineering
University of Southern California
Los Angeles, California

John P. Uyemura
School of Electrical Engineering
Georgia Institute of Technology
Atlanta, Georgia

Chorng-Kuang Wang
Department of Electrical Engineering
National Taiwan University
Taipei, Taiwan

Bogdan M. Wilamowski
Alabama Nano/Micro Science and Technology Center
Department of Electrical and Computer Engineering
Auburn University
Auburn, Alabama

Kung Yao
Electrical Engineering Department
University of Southern California, Los Angeles
Los Angeles, California

1

Monolithic Device Models

1.1 Bipolar Junction Transistor **1**-1
Ebers–Moll Model • Gummel–Poon Model • Current Gains of Bipolar Transistors • High-Current Phenomena • Small-Signal Model • Technologies • Model Parameters • SiGe HBTs
References **1**-20
1.2 Metal–Oxide–Silicon Field Effect Transistor **1**-21
Introduction • Channel Charge • Volt–Ampere Characteristics • Transistor Capacitances • Small-Signal Operation • Design-Oriented Analysis Strategy
References **1**-81
1.3 JFET, MESFET, and HEMT Technology and Devices **1**-82
Introduction • Silicon JFET Device Operation and Technology • Compound Semiconductor FET Technologies • Conclusion
References **1**-101
1.4 Passive Components **1**-103
Resistors • Capacitors • Inductors
References **1**-131
1.5 Chip Parasitics in Analog Integrated Circuits **1**-132
Interconnect Parasitics • Pad and Packaging Parasitics • Parasitic Measurement
References **1**-145

Bogdan M. Wilamowski
Auburn University

Guofu Niu
Auburn University

John Choma, Jr.
University of Southern California

Stephen I. Long
University of California, Santa Barbara

Nhat M. Nguyen
Rambus Inc.

Martin A. Brooke
Georgia Institute of Technology

1.1 Bipolar Junction Transistor

Bogdan M. Wilamowski and Guofu Niu

The bipolar junction transistor (BJT) is historically the first solid-state analog amplifier and digital switch, and formed the basis of integrated circuits (ICs) in the 1970s. Starting in the early 1980s, the MOSFET had gradually taken over, particularly for main stream digital ICs. However, in the 1990s, the invention of silicon–germanium base heterojunction bipolar transistor (SiGe HBT) brought the bipolar transistor back into high-volume commercial production, mainly for the now widespread wireless and wire line communications applications. Today, SiGe HBTs are used to design radio-frequency (RF) ICs and systems for cell phones, wireless local area network (WLAN), automobile collision avoidance

radar, wireless distribution of cable television, millimeter wave radios, and many more applications, due to its outstanding high-frequency performance and ability to integrate with CMOS for realizing digital, analog, and RF functions on the same chip.

Below we first introduce the basic concepts of BJT using a historically important equivalent circuit model, the Ebers–Moll model. Then the Gummel–Poon model is introduced, as it is widely used for computer-aided design, and is the basis of modern BJT models like the VBIC, Mextram, and HICUM models. Current gain, high-current phenomena, fabrication technologies, and SiGe HBTs are then discussed.

1.1.1 Ebers–Moll Model

A NPN BJT consists of two closely spaced PN junctions connected back to back sharing the same p-type region, as shown in Figure 1.1a. The drawing is not drawn to scale. The emitter and base layers are thin, typically less than 1 μm, and the collector is much thicker to support a high output voltage swing. For forward mode operation, the emitter–base (EB) junction is forward biased, and the collector–base (CB) junction is reverse biased. Minority carriers are injected from emitter to base, travel across the base, and are then collected by the reverse biased CB junction. Therefore, the collector current is transported from the EB junction, and thus proportional to the EB junction current. In the forward-active mode, the current–voltage characteristic of the EB junction is described by the well-known diode equation

$$I_{EF} = I_{E0}\left[\exp\left(\frac{V_{BE}}{V_T}\right) - 1\right] \tag{1.1}$$

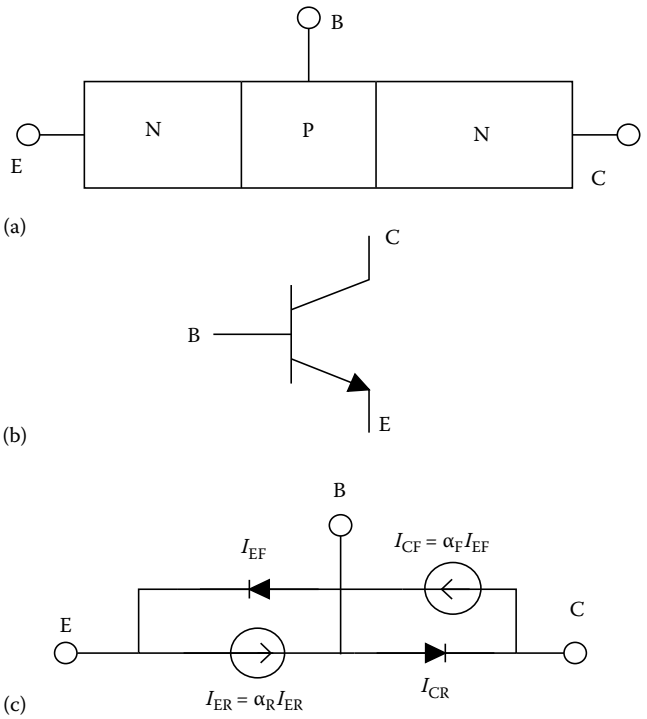


FIGURE 1.1 (a) Cross-sectional view of a NPN BJT. (b) Circuit symbol. (c) The Ebers–Moll equivalent circuit model.

where

I_{E0} is the EB junction saturation current

$V_T = kT/q$ is the thermal potential (about 25 mV at room temperature)

The collector current is typically smaller than the emitter current $I_{CF} = \alpha_F I_{EF}$, where α_F is the forward current gain.

Under reverse mode operation, the CB junction is forward biased and the EB junction is reverse biased. Like in the forward mode, the forward biased CB junction current gives the collector current

$$I_{CF} = I_{C0}\left[\exp\left(\frac{V_{BC}}{V_T}\right) - 1\right] \tag{1.2}$$

where I_{C0} is the CB junction saturation current. Similarly $I_{ER} = \alpha_R I_R$, where α_R is the reverse current gain. Under general biasing conditions, it can be proven that to first order, a superposition of the above described forward and reverse mode equivalent circuits can be used to describe transistor operation, as shown in Figure 1.1b. The forward transistor operation is described by Equation 1.1, and the reverse transistor operation is described by Equation 1.2. From the Kirchoff's current law one can write $I_C = I_{CF} - I_{CR}$, $I_E = I_{EF} - I_{ER}$, and $I_B = I_E - I_C$. Using Equations 1.1 and 1.2 the emitter and collector currents can be described as

$$\begin{aligned} I_E &= a_{11}\left(\exp\frac{V_{BE}}{V_T} - 1\right) - a_{12}\left(\exp\frac{V_{BC}}{V_T} - 1\right) \\ I_C &= a_{21}\left(\exp\frac{V_{BE}}{V_T} - 1\right) - a_{22}\left(\exp\frac{V_{BC}}{V_T} - 1\right) \end{aligned} \tag{1.3}$$

which are known as the Ebers–Moll equations [1]. The Ebers–Moll coefficients a_{ij} are given as

$$a_{11} = I_{E0}, \quad a_{12} = \alpha_R I_{C0}, \quad a_{21} = \alpha_F I_{E0}, \quad a_{22} = I_{C0} \tag{1.4}$$

The Ebers–Moll coefficients are a very strong function of the temperature

$$a_{ij} = K_x T^m \exp\frac{V_{go}}{V_T} \tag{1.5}$$

where

K_x is proportional to the junction area and independent of the temperature

$V_{go} = 1.21$ V is the bandgap voltage in silicon (extrapolated to 0 K)

m is a material constant with a value between 2.5 and 4

When both EB and CB junctions are forward biased, the transistor is called to be working in the saturation region. Current injection through the collector junction may activate parasitic transistors in ICs using p-type substrate, where base acts as emitter, collector as base, and substrate as collector. In typical ICs, bipolar transistors must not operate in saturation. Therefore, for the integrated bipolar transistor the Ebers–Moll equations can be simplified to the form

$$\begin{aligned} I_E &= a_{11}\left(\exp\frac{V_{BE}}{V_T} - 1\right) \\ I_C &= a_{21}\left(\exp\frac{V_{BE}}{V_T} - 1\right) \end{aligned} \tag{1.6}$$

where $a_{21}/a_{11} = \alpha_F$. This equation corresponds to the circuit diagram shown in Figure 1.1c.

1.1.2 Gummel–Poon Model

In real bipolar transistors the current voltage characteristics are more complex than those described by the Ebers–Moll equations. Typical current–voltage characteristics of the bipolar transistor, plotted in semilogarithmic scale, are shown in Figure 1.2. At small-base emitter voltages, due to the generation–recombination phenomena, the base current is proportional to

$$I_{BL} \propto \exp\frac{V_{BE}}{2V_T} \tag{1.7}$$

Also, due to the base conductivity modulation at high-level injections, the collector current for larger voltages can be expressed by the similar relation

$$I_{CH} \propto \exp\frac{V_{BE}}{2V_T} \tag{1.8}$$

Note, that the collector current for wide range is given by

$$I_C = I_s \exp\frac{V_{BE}}{V_T} \tag{1.9}$$

The saturation current is a function of device structure parameters

$$I_s = \frac{qAn_i^2 V_T \mu_B}{\int_0^{w_B} N_B(x)\mathrm{d}x} \tag{1.10}$$

where

$q = 1.6 \times 10^{-19}$ C is the electron charge
A is the EB junction area
n_i is the intrinsic concentration ($n_i = 1.5 \times 10^{10}$ at 300 K)
μ_B is the mobility of the majority carriers in the transistor base
w_B is the effective base thickness
$N_B(x)$ is the distribution of impurities in the base

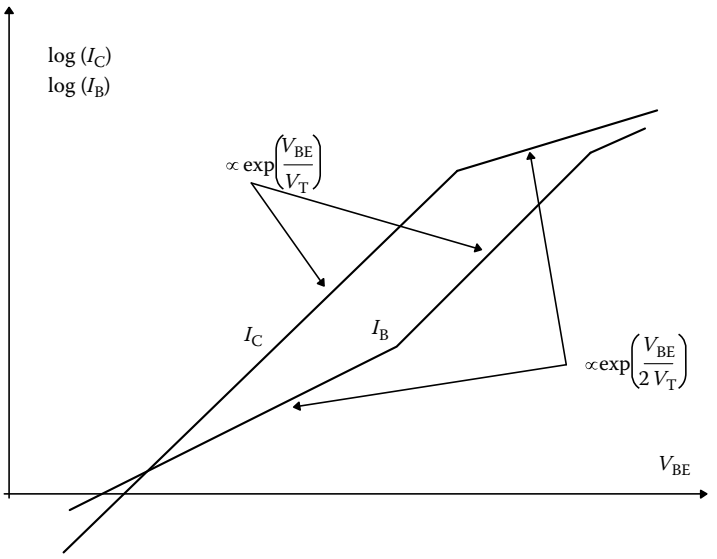


FIGURE 1.2 Collector and base currents as a function of base–emitter voltage.

Note, that the saturation current is inversely proportional to the total impurity dose in the base. In the transistor with the uniform base, the saturation current is given by

$$I_s = \frac{qAn_i^2 V_T \mu_B}{w_B N_B} \tag{1.11}$$

When a transistor operates in the reverse-active mode (emitter and collector are switched) then the current of such biased transistor is given by

$$I_E = I_s \exp \frac{V_{BC}}{V_T} \tag{1.12}$$

Note, that the I_s parameter is the same for forward and reverse mode of operation. The Gummel–Poon transistor model [2] was derived from the Ebers–Moll model using the assumption that $a_{12} = a_{21} = I_s$. For the Gummel–Poon model, Equations 1.3 are simplified to the form

$$\begin{aligned} I_E &= I_s \left(\frac{1}{\alpha_F} \exp \frac{V_{BE}}{V_T} - \exp \frac{V_{BC}}{V_T} \right) \\ I_C &= I_s \left(\exp \frac{V_{BE}}{V_T} - \frac{1}{\alpha_R} \exp \frac{V_{BC}}{V_T} \right) \end{aligned} \tag{1.13}$$

These equations require only three coefficients, while the Ebers–Moll requires four. The saturation current I_s is constant for a wide range of currents. The current gain coefficients α_F and α_R have values smaller, but close to unity. Often instead of using the current gain as $\alpha = I_C/I_E$, the current gain β as a ratio of the collector current to the base current $\beta = I_C/I_B$ is used. The mutual relationships between α and β coefficients are given by

$$\alpha_F = \frac{\beta_F}{\beta_F + 1}, \quad \beta_F = \frac{\alpha_F}{1 - \alpha_F}, \quad \alpha_R = \frac{\beta_R}{\beta_R + 1}, \quad \beta_R = \frac{\alpha_R}{1 - \alpha_R} \tag{1.14}$$

The Gummel–Poon model was implemented in Simulation Program with Integrated Circuit Emphasis (SPICE) [3] and other computer programs for circuit analysis. To make the equations more general, the material parameters η_F and η_R were introduced

$$I_C = I_s \left[\exp \frac{V_{BE}}{\eta_F V_T} - \left(1 + \frac{1}{\beta_R} \right) \exp \frac{V_{BC}}{\eta_R V_T} \right] \tag{1.15}$$

The values of η_F and η_R vary from 1 to 2.

1.1.3 Current Gains of Bipolar Transistors

The transistor current gain β is limited by two phenomena: base transport efficiency and emitter injection efficiency. The effective current gain β can be expressed as

$$\frac{1}{\beta} = \frac{1}{\beta_I} + \frac{1}{\beta_T} + \frac{1}{\beta_R} \tag{1.16}$$

where

β_I is the transistor current gain caused by emitter injection efficiency
β_T is the transistor current gain caused by base transport efficiency
β_R is the recombination component of the current gain

As one can see from Equation 1.16, smaller values of β_I, β_T, and β_R dominate. The base transport efficiency can be defined as a ratio of injected carriers into the base, to the carriers that recombine within the base. This ratio is also equal to the ratio of the minority carrier life time, to the transit time of carriers through the base. The carrier transit time can be approximated by an empirical relationship

$$\tau_{transit} = \frac{w_B^2}{V_T \mu_B (2 + 0.9\eta)}, \quad \eta = \ln\left(\frac{N_{BE}}{N_{BC}}\right) \tag{1.17}$$

where
- μ_B is the mobility of the minority carriers in base
- w_B is the base thickness
- N_{BE} is the impurity doping level at the emitter side of the base
- N_{BC} is the impurity doping level at the collector side of the base

Therefore, the current gain due to the transport efficiency is

$$\beta_T = \frac{\tau_{life}}{\tau_{transit}} = (2 + 0.9\eta)\left(\frac{L_B}{w_B}\right)^2 \tag{1.18}$$

where $L_B = \sqrt{V_T \mu_B \tau_{life}}$ is the diffusion length of minority carriers in the base.

The current gain β_I, due to the emitter injection efficiency, is given

$$\beta_I = \frac{\mu_B \int_0^{w_E} N_{Eeff}(x)dx}{\mu_E \int_0^{w_B} N_B(x)dx} \tag{1.19}$$

where
- μ_B and μ_E are minority carrier mobilities in the base and in the emitter
- $N_B(x)$ is impurity distribution in the base
- N_{Eeff} is the effective impurity distribution in the emitter

The recombination component of current gain β_R is caused by the different current–voltage relationship of base and collector currents as can be seen in Figure 1.2. The slower base current increase is due to the recombination phenomenon within the depletion layer of the base–emitter junction. Since the current gain is a ratio of the collector current to the base current, the relation for β_R can be found as

$$\beta_R = K_{R0} I_C^{1-(1/\eta_R)} \tag{1.20}$$

As it can be seen from Figure 1.2, the current gain β is a function of the current. This gain–current relationship is illustrated in Figure 1.3. The range of a constant current gain is wide for bipolar transistors with a technology characterized by a lower number of generation–recombination centers.

With an increase of CB voltage, the depletion layer penetrates deeper into the base. Therefore, the effective thickness of the base decreases. This leads to an increase of transistor current gain with applied collector voltages. Figure 1.4 illustrates this phenomenon, which is known as the Early's effect. The extensions of transistor characteristics (dotted lines in Figure 1.4) are crossing the voltage axis at

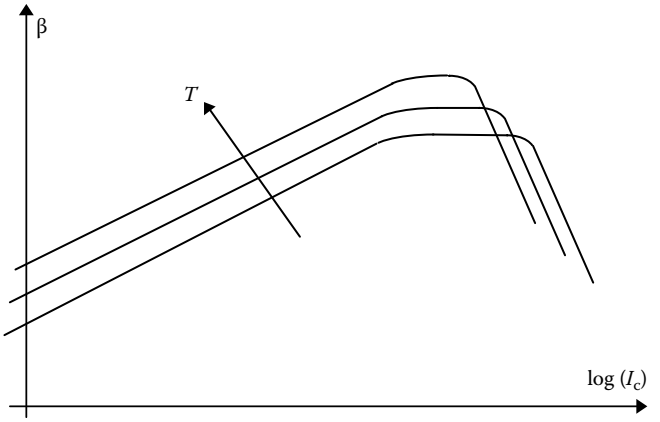


FIGURE 1.3 Current gain β as a function of collector current.

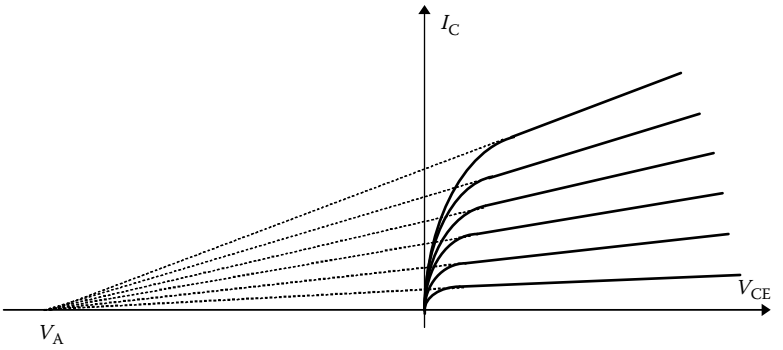


FIGURE 1.4 Current–voltage characteristics of a bipolar transistor.

the point $-V_A$, where V_A is known as the Early voltage. The current gain β, as a function of collector voltage, is usually expressed using the relation

$$\beta = \beta_o\left(1 + \frac{V_{CE}}{V_A}\right) \tag{1.21}$$

Similar equation can be defined for the reverse mode of operation.

1.1.4 High-Current Phenomena

The concentration of minority carriers increases with the rise of transistor currents. When the concentration of moving carriers exceeds a certain limit, the transistor property degenerates. Two phenomena are responsible for this limitation. The first is related to the high concentration of moving carriers (electrons in the NPN transistor) in the base–collector depletion region. This is known as the Kirk effect. The second phenomenon is caused by a high level of carriers injected into the base. When the concentration of injected minority carriers in the base exceeds the impurity concentration there, then the base conductivity modulation limits the transistor performance.

To understand the Kirk effect consider the NPN transistor in forward-active mode with the base–collector junction reversely biased. The depletion layer consists of the negative lattice charge of the base

region and the positive lattice charge of the collector region. Boundaries of the depletion layer are such that total the positive and negative charges are equal. When a collector current, carrying negatively charged electrons, flows through the junction, effective negative charge on the base side of junction increases. Also, the positive lattice charge of the collector side of the junction is compensated by negative charge of moving electrons. This way, the CB space charge region moves toward the collector, resulting in a thicker effective base. With a large current level, the thickness of the base may be doubled or tripled. This phenomenon, known as the Kirk effect, becomes very significant when the charge of moving electrons exceeds the charge of the lightly doped collector N_C. The threshold current for the Kirk effect is given by

$$I_{max} = qAv_{sat}N_C \tag{1.22}$$

where v_{sat} is the saturation velocity for electrons ($v_{sat} = 10^7$ cm/s for silicon).

The conductivity modulation in the base, or high-level injection, starts when the concentration of injected electrons into the base exceeds the lowest impurity concentration in the base N_{Bmin}. This occurs for the collector current I_{max} given by

$$I_{max} < qAN_{Bmax}, \quad \nu = \frac{qAV_T\mu_B N_{Bmax}(2 + 0.9\eta)}{w_B} \tag{1.23}$$

The above equation is derived using Equation 1.17 for the estimation of base transient time.

The high-current phenomena are significantly enlarged by the current crowding effect. The typical cross section of bipolar transistor is shown in Figure 1.5. The horizontal flow of the base current results in the voltage drop across the base region under the emitter. This small voltage difference on the base–emitter junction causes a significant difference in the current densities at the junction. This is due to the very nonlinear junction current–voltage characteristics. As a result, the base–emitter junction has very nonuniform current distribution across the junction. Most of the current flows through the part of the junction closest to base contact. For transistors with larger emitter areas, the current crowding effect is more significant. This nonuniform transistor current distribution makes the high-current phenomena, such as the base conductivity modulation and the Kirk effect, start for smaller currents than given by Equations 1.22 and 1.23. The current crowding effect is also responsible for the change of the effective base resistance with a current. As base current increases, the larger part of emitter current flows closer to the base contact, and the effective base resistance decreases.

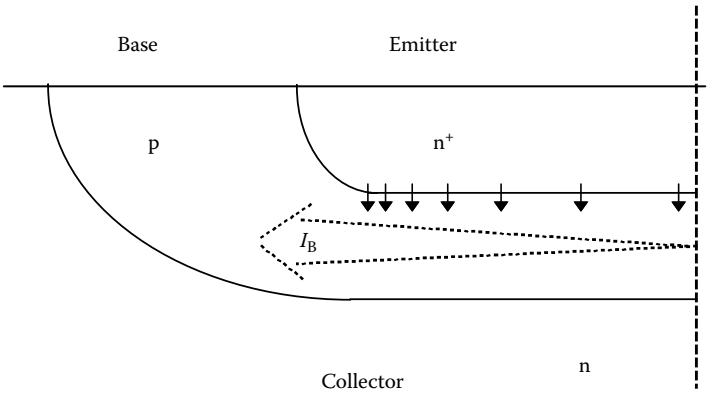


FIGURE 1.5 Current crowding effect.

1.1.5 Small-Signal Model

Small-signal transistor models are essential for AC circuit design. The small-signal equivalent circuit of the bipolar transistor is shown in Figure 1.6a. The lumped circuit shown in Figure 1.6a is only an approximation. In real transistors resistances and capacitances have a distributed character. For most design tasks, this lumped model is adequate, or even the simple equivalent transistor model shown in Figure 1.6b can be considered. The small-signal resistances, r_π and r_o, are inversely proportional to the transistor currents, and the transconductance g_m is directly proportional to the transistor currents

$$r_\pi = \frac{\eta_F V_T}{I_B} = \frac{\eta_F V_T \beta_F}{I_C}, \quad r_o = \frac{V_A}{I_C}, \quad g_m = \frac{I_C}{\eta_F V_T} \tag{1.24}$$

where

η_F is the forward emission coefficient, ranging form 1.0 to 2.0
V_T is the thermal potential ($V_T = 25$ mV at room temperature)

Similar equations to Equation 1.24 can be written for the reverse transistor operation as well.

The series base, emitter, and collector resistances R_B, R_E, and R_C are usually neglected for simple analysis (Figure 1.6b). However, for high-frequency analysis it is essential to use at least the base series resistance R_B. The series emitter resistance R_E usually has a constant, bias-independent value. The collector resistance R_C may significantly vary with the biasing current. The value of the series collector resistance may lower by one or two orders of magnitude if the collector junction becomes forward biased. A large series collector resistance may force the transistor into the saturation mode. Usually, when collector–emitter voltage is large enough, the effect of collector resistance is not significant. The SPICE model assumes constant value for the collector resistance R_C.

The series base resistance R_B may significantly limit the transistor performance at high frequencies. Due to the current crowding effect and the base conductivity modulation, the series base resistance is a function of the collector current I_C [4]

$$R_B = R_{Bmin} + \frac{R_{B0} - R_{Bmin}}{0.5 + \sqrt{0.25 + \frac{i_C}{I_{KF}}}} \tag{1.25}$$

where

I_{KF} is β_F high-current roll-off corner
R_{B0} is the base resistance at very small currents
R_{Bmin} is the minimum base resistance at high currents

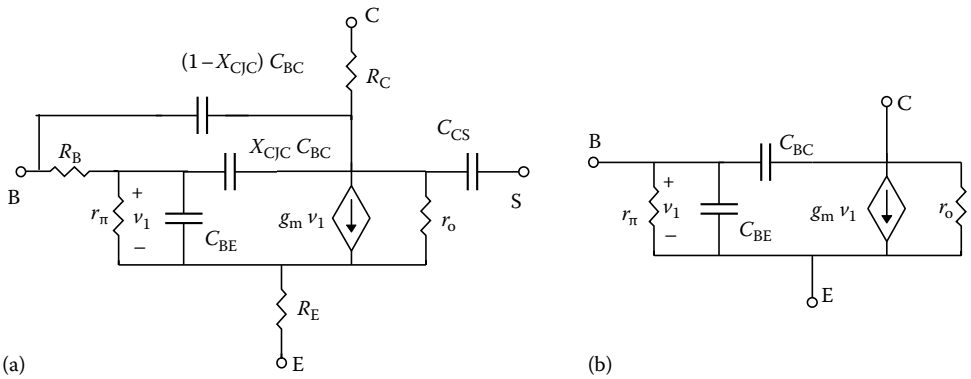


FIGURE 1.6 Bipolar transistor equivalent diagrams. (a) SPICE model and (b) simplified model.

Another possible approximation of the base series resistance R_B, as a function of the base current I_B, is [4]

$$R_B = 3(R_{B0} - R_{Bmin})\frac{\tan z - z}{z\tan^2 z} + R_{Bmin}, \quad z = \frac{\sqrt{1 + \frac{1.44 I_B}{\pi^2 I_{RB}}} - 1}{\frac{24}{\pi^2}\sqrt{\frac{I_B}{I_{RB}}}} \tag{1.26}$$

where I_{RB} is the base current for which the base resistance falls halfway to its minimum value.

The base–emitter capacitance C_{BE} is composed of two terms: the diffusion capacitance, which is proportional to the collector current, and the depletion capacitance, which is a function of the base–emitter voltage V_{BE}. The C_{BE} capacitance is given by

$$C_{BE} = \tau_F \frac{I_C}{\eta_F V_T} + C_{JE0}\left(1 - \frac{v_{BE}}{V_{JE0}}\right)^{-m_{JE}} \tag{1.27}$$

where

V_{JE0} is the base–emitter junction potential
τ_F is the base transit time for forward direction
C_{JE0} is the base–emitter zero-bias junction capacitance
m_{JE} is the base–emitter grading coefficient

The base–collector capacitance C_{BC} is given by a similar expression as Equation 1.27. In the case when the transistor operates in forward-active mode, it can be simplified to

$$C_{BC} = C_{JC0}\left(1 - \frac{v_{BC}}{V_{JC0}}\right)^{-m_{JC}} \tag{1.28}$$

where

V_{JC0} is the base–collector junction potential
C_{JC0} is the base–collector zero-bias junction capacitance
m_{JC} is the base–collector grading coefficient

In the case when the bipolar transistor is in the integrated form, the collector–substrate capacitance C_{CS} has to be considered

$$C_{CS} = C_{JS0}\left(1 - \frac{V_{CS}}{V_{JS0}}\right)^{-m_{JS}} \tag{1.29}$$

where

V_{JS0} is the collector–substrate junction potential
C_{JS0} the collector–substrate zero-bias junction capacitance
m_{JS} is the collector–substrate grading coefficient

When the transistor enters saturation, or it operates in the reverse-active mode, Equations 1.27 and 1.28 should be modified to

$$C_{BE} = \tau_F \frac{I_S \exp\left(\frac{V_{BE}}{\eta_F V_T}\right)}{\eta_F V_T} + C_{JE0}\left(1 - \frac{V_{BE}}{V_{JE0}}\right)^{-m_{JE}} \tag{1.30}$$

$$C_{BC} = \tau_R \frac{I_S \exp\left(\frac{V_{BC}}{\eta_R V_T}\right)}{\eta_R V_T} + C_{JC0}\left(1 - \frac{V_{BC}}{V_{JC0}}\right)^{-m_{JC}} \tag{1.31}$$

1.1.6 Technologies

The bipolar technology was used to fabricate the first ICs more than 40 years ago. A similar standard bipolar process is still used. In recent years, for high-performance circuits and for BiCMOS technology, the standard bipolar process was modified by using the thick selective silicon oxidation instead of the p-type isolation diffusion. Also, the diffusion process was substituted by the ion implantation process, low-temperature epitaxy, and Chemical Vapor Deposition (CVD).

1.1.6.1 Integrated NPN Bipolar Transistor

The structure of the typical integrated bipolar transistor is shown in Figure 1.7. The typical impurity profile of the bipolar transistor is shown in Figure 1.8. The emitter doping level is much higher than the base doping, so large current gains are possible (see Equation 1.19). The base is narrow and it has an impurity gradient, so the carrier transit time through the base is short (see Equation 1.17). Collector concentration near the base–collector junction is low, therefore, the transistor has a large breakdown voltage, large Early voltage V_{AF}, and CB depletion capacitance is low. High impurity concentration in the buried layer leads to a small collector series resistance. The emitter strips have to be as narrow as technology allows, reducing the base series resistance and the current crowding effect. If large emitter area is required, many narrow emitter strips interlaced with base contacts have to be used in a single

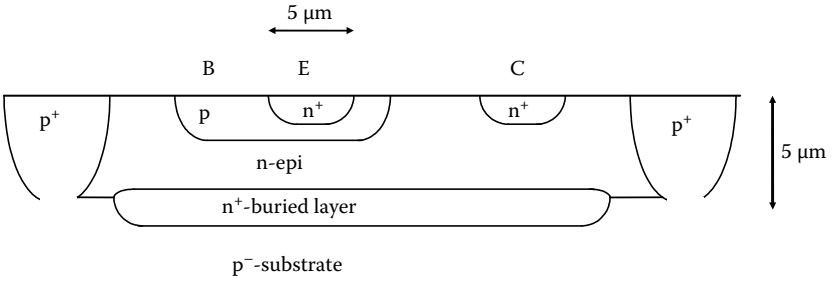


FIGURE 1.7 NPN bipolar structure.

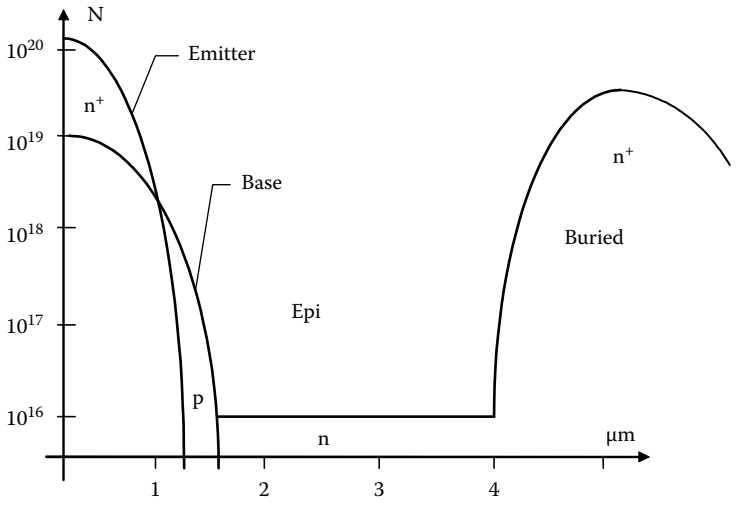


FIGURE 1.8 Cross section of a typical bipolar transistor.

transistor. Special attention has to be taken during the circuit design, so the base–collector junction is not forward biased. If the base–collector junction is forward biased, then the parasitic PNP transistors activate. This leads to undesired circuit operation. Thus, the integrated bipolar transistors must not operate in reverse or in saturation modes.

1.1.6.2 Lateral and Vertical PNP Transistors

The standard bipolar technology is oriented for fabrication of the NPN transistors with the structure shown in Figure 1.7. Using the same process, other circuit elements, such as resistors and PNP transistors, can be fabricated as well.

The lateral transistor, shown in Figure 1.9a uses the base p-type layer for both emitter and collector fabrication. The vertical transistor, shown in Figure 1.9b uses the p-type base layer for emitter, and the p-type substrate as collector. This transistor is sometimes known as the substrate transistor. In both transistors the base is made of the n-type epitaxial layer. Such transistors with a uniform and thick base are slow. Also, the current gain β of such transistors is small. Note, that the vertical transistor has the collector shorted to the substrate as Figure 1.10b illustrates. When a PNP transistor with a large current gain is required, then the concept of the composite transistor can be implemented. Such a composite transistor, known also as superbeta transistor, consists a PNP lateral transistor, and the standard NPN transistor connected as shown in Figure 1.10c. The composed transistor acts as the PNP transistor and it has a current gain β approximately equal to $\beta_{pnp}\beta_{npn}$.

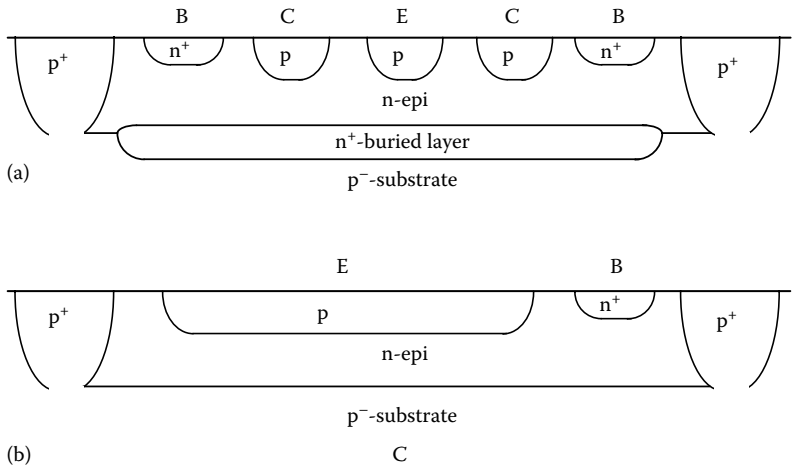


FIGURE 1.9 Integrated PNP transistors: (a) lateral PNP transistor, and (b) substrate PNP transistor.

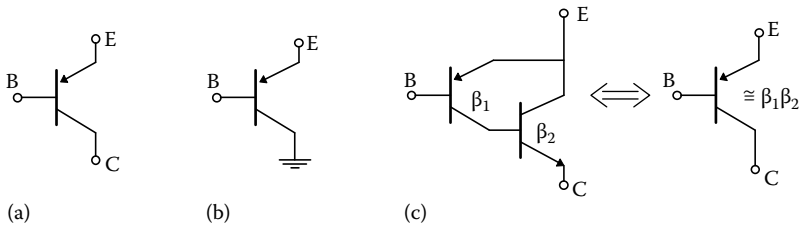


FIGURE 1.10 Integrated PNP transistors: (a) lateral transistor, (b) substrate transistor, and (c) composed transistor.

1.1.7 Model Parameters

It is essential to use proper transistor models in the computer aided design tools. The accuracy of simulation results depends on the model accuracy, and on the values of the model parameters used. In Section 1.1, the thermal and second-order effect in the transistor model are discussed. The SPICE bipolar transistor model parameters are discussed.

1.1.7.1 Thermal Sensitivity

All parameters of the transistor model are temperature dependent. Some parameters are very strong functions of temperature. To simplify the model description, the temperature dependence of some parameters are often neglected. In this chapter, the temperature dependence of the transistor model is described based on the model of the SPICE program [3–5]. Deviations from the actual temperature dependence will also be discussed. The temperature dependence of junction capacitance is given by

$$C_J(T) = C_J\left\{1 + m_J\left[4.010^{-4}(T - T_{NOM}) + 1 - \frac{V_J(T)}{V_J}\right]\right\} \tag{1.32}$$

where T_{NOM} is the nominal temperature, which is specified in the SPICE program in the .OPTIONS statement. The junction potential $V_J(T)$ is a function of temperature

$$V_J(T) = V_J\frac{T}{T_{NOM}} - 3V_T\ln\left(\frac{T}{T_{NOM}}\right) - E_G(T) + E_G\frac{T}{T_{NOM}} \tag{1.33}$$

The value of 3 in the multiplication coefficient of above equation is from the temperature dependence of the effective state densities in the valence and conduction bands. The temperature dependence of the energy gap is computed in the SPICE program from

$$E_G(T) = E_G - \frac{7.0210^{-4}T^2}{T + 1108} \tag{1.34}$$

The transistor saturation current as a function of temperature is calculated as

$$I_s(T) = I_s\left(\frac{T}{T_{NOM}}\right)^{X_{TI}}\exp\left[\frac{E_G(T - T_{NOM})}{V_T T_{NOM}}\right] \tag{1.35}$$

where E_G is the energy gap at the nominal temperature. The junction leakage currents I_{SE} and I_{SC} are calculated using

$$I_{SE}(T) = I_{SE}\left(\frac{T}{T_{NOM}}\right)^{X_{TI}-X_{TB}}\exp\left[\frac{E_G(T - T_{NOM})}{\eta_E V_T T_{NOM}}\right] \tag{1.36}$$

and

$$I_{SC}(T) = I_{SC}\left(\frac{T}{T_{NOM}}\right)^{X_{TI}-X_{TB}}\exp\left[\frac{E_G(T - T_{NOM})}{\eta_C V_T T_{NOM}}\right] \tag{1.37}$$

The temperature dependence of the transistor current gains β_F and β_R are modeled in the SPICE as

$$\beta_F(T) = \beta_F\left(\frac{T}{T_{NOM}}\right)^{X_{TB}}, \quad \beta_R(T) = \beta_R\left(\frac{T}{T_{NOM}}\right)^{X_{TB}} \tag{1.38}$$

The SPICE model does not give accurate results for the temperature relationship of the current gain β at high currents. For high current levels the current gain decreases sharply with the temperature, as can be seen from Figure 1.3. Also, the knee current parameters IKF, IKR, IKB are temperature-dependent, and this is not implemented in the SPICE program.

1.1.7.2 Second-Order Effects

The current gain β is sometimes modeled indirectly by using different equations for the collector and base currents [4,5]

$$I_C = \frac{I_S(T)}{Q_b}\left(\exp\frac{V_{BE}}{\eta_F V_T} - \exp\frac{V_{BC}}{\eta_R V_T}\right) - \frac{I_S(T)}{\beta_R(T)}\left(\exp\frac{V_{BC}}{\eta_R V_T} - 1\right) - I_{SC}(T)\left(\exp\frac{V_{BC}}{\eta_C V_T} - 1\right) \tag{1.39}$$

where

$$Q_b = \frac{1 + \sqrt{1 + 4Q_X}}{2\left(1 - \frac{V_{BC}}{V_{AF}} - \frac{V_{BE}}{V_{AR}}\right)} \tag{1.40}$$

$$Q_X = \frac{I_s(T)}{I_{KF}}\left(\exp\frac{V_{BE}}{\eta_F V_T} - 1\right) + \frac{I_s(T)}{I_{KR}}\left(\exp\frac{V_{BC}}{\eta_R V_T} - 1\right) \tag{1.41}$$

and

$$I_B = \frac{I_S}{\beta_F}\left(\exp\frac{V_{BE}}{\eta_F V_T} - 1\right) + I_{SE}\left(\exp\frac{V_{BE}}{\eta_E V_T} - 1\right) + \frac{I_S}{\beta_R}\left(\exp\frac{V_{BC}}{\eta_R V_T} - 1\right) + I_{SC}\left(\exp\frac{V_{BC}}{\eta_C V_T} - 1\right) \tag{1.42}$$

where

I_{SE} is the base–emitter junction leakage current
I_{SC} is the base–collector junction leakage current
η_E is the base–emitter junction leakage emission coefficient
η_C is the base–collector junction leakage emission coefficient

The forward transit time τ_F is a function of biasing conditions. In the SPICE program the τ_F parameter is computed using

$$\tau_F = \tau_{F0}\left[1 + X_{TF}\left(\frac{I_{CC}}{I_{CC} + I_{TF}}\right)^2 \exp\frac{V_{BC}}{1.44 V_{TF}}\right], \quad I_{CC} = I_s\left(\exp\frac{V_{BE}}{\eta_F V_T} - 1\right) \tag{1.43}$$

At high frequencies the phase of the collector current shifts. This phase shift is computed in the SPICE program following way

$$I_C(\omega) = I_C \exp(j\omega P_{TF}\tau_F) \tag{1.44}$$

where P_{TF} is a coefficient for excess phase calculation.

Noise is usually modeled as the thermal noise for parasitic series resistances, and as shot and flicker noise for collector and base currents

$$\overline{i_B^2} = \frac{4kT\Delta f}{R} \tag{1.45}$$

$$\overline{i_B^2} = \left(2qI_B + \frac{K_F I_B^{A_F}}{F}\right)\Delta f \tag{1.46}$$

$$\overline{i_C^2} = 2qI_C\Delta f \tag{1.47}$$

where K_F and A_F are the flicker-noise coefficient and flicker-noise exponent. More detailed information about noise modeling is given in Section 3.2.

1.1.7.3 SPICE Model of the Bipolar Transistor

The SPICE model of bipolar transistor uses similar or identical equations as described in this chapter [3–5]. Table 1.1 shows the parameters of the bipolar transistor model and its relation to the parameters used in this chapter.

TABLE 1.1 Parameters of SPICE Bipolar Transistor Model

Name Used	Equations	SPICE Name	Parameter Description	Unit	Typical Value	SPICE Default
I_s	1.10, 1.11	IS	Saturation current	A	10^{-15}	10^{-16}
I_{SE}	1.39	ISE	B–E leakage saturation current	A	10^{-12}	0
I_{SC}	1.39	ICS	B–C leakage saturation current	A	10^{-12}	0
β_F	1.14, 1.16, 1.21	BF	Forward current gain	—	100	100
β_R	1.14, 1.16, 1.21	BF	Reverse current gain	—	0.1	1
η_F	1.15, 1.24, 1.30, 1.31, 1.39 through 1.41	NF	Forward current emission coefficient	—	1.2	1.0
η_R	1.15, 1.24, 1.30, 1.31, 1.39 through 1.42	NR	Reverse current emission coefficient	—	1.3	1.0
η_E	1.39	NE	B–E leakage emission coefficient	—	1.4	1.5
η_C	1.39	NC	B–C leakage emission coefficient	—	1.4	1.5
V_{AF}	1.21, 1.40	VAF	Forward Early voltage	V	200	∞
V_{AR}	1.21, 1.40	VAR	Reverse Early voltage	V	50	∞
I_{KF}	1.22, 1.23, 1.40	IKF	β_F high-current roll-off corner	A	0.05	∞
I_{KR}	1.22, 1.23, 1.40	IKR	β_R high-current roll-off corner	A	0.01	∞
I_{RB}	1.26	IRB	Current where base resistance falls by half	A	0.1	∞
R_B	1.25, 1.26	RB	Zero base resistance	Ω	100	0
R_{Bmin}	1.25, 1.26	RBM	Minimum base resistance	Ω	10	RB
R_E	Figure 1.6	RE	Emitter series resistance	Ω	1	0
R_C	Figure 1.6	RC	Collector series resistance	Ω	50	0
C_{JE0}	1.27	CJE	B–E zero-bias depletion capacitance	F	10^{-12}	0
C_{JC0}	1.28	CJC	B–C zero-bias depletion capacitance	F	10^{-12}	0
C_{JS0}	1.29	CJS	Zero-bias collector–substrate capacitance	F	10^{-12}	0
V_{JE0}	1.27	VJE	B–E built-in potential	V	0.8	0.75
V_{JC0}	1.28	VJC	B–C built-in potential	V	0.7	0.75
V_{JS0}	1.29	VJS	Substrate junction built-in potential	V	0.7	0.75
m_{JE}	1.27	MJE	B–E junction exponential factor	—	0.33	0.33
m_{JC}	1.28	MJC	B–C junction exponential factor	—	0.5	0.33

(*continued*)

TABLE 1.1 (continued) Parameters of SPICE Bipolar Transistor Model

Name Used	Equations	SPICE Name	Parameter Description	Unit	Typical Value	SPICE Default
m_{JS}	1.29	MJS	Substrate junction exponential factor	—	0.5	0
X_{CJC}	Figure 1.6	XCJC	Fraction of B–C capacitance connected to internal base node (see Figure 1.6)	—	0.5	0
τ_F	1.17, 1.28, 1.30, 1.42	TF	Ideal forward transit time	s	10^{-10}	0
τ_R	1.31	TR	Reverse transit time	s	10^{-8}	0
X_{TF}	1.43	XTF	Coefficient for bias dependence of τ_F	—		0
V_{TF}	1.43	VTF	Voltage for τ_F dependence on V_{BC}	V		∞
I_{TF}	1.43	ITF	Current where $\tau_F = f(I_C, V_{BC})$ starts	A		0
P_{TF}	1.44	PTF	Excess phase at freq $= 1/(2\pi\tau_F)$ Hz	°		0
X_{TB}	1.38	XTB	Forward and reverse beta temperature exponent			0
E_G	1.34	EG	Energy gap	eV	1.1	1.11
X_{TI}	1.35 through 1.37	XTI	Temperature exponent for effect on I_s	—	3.5	3
K_F	1.46	KF	Flicker-noise coefficient	—		0
A_F	1.46	AF	Flicker-noise exponent	—		1
F_C		FC	Coefficient for the forward biased depletion capacitance formula	—	0.5	0.5
T_{NOM}	1.32 through 1.38	TNOM	Nominal temperature specified in .OPTION statement	K	300	300

The SPICE [3] was developed mainly for analysis of ICs. During the analysis it is assumed that the temperatures of all circuit elements are the same. This is not true for power ICs where the junction temperatures may differ by 30 K or more. This is obviously not true for circuits composed of the discrete elements where the junction temperatures may differ by 100 K and more. These temperature effects, which can significantly affect the analysis results, are not implemented in the SPICE program.

Although the SPICE bipolar transistor model uses more than 40 parameters, many features of the bipolar transistor are not included in the model. For example, the reverse junction characteristics are described by Equation 1.32. This model does not give accurate results. In the real silicon junction the leakage current is proportional to the thickness of the depletion layer, which is proportional to $V^{1/m}$. Also the SPICE model of the bipolar transistor assumes that there is no junction breakdown voltage. A more accurate model of the reverse junction characteristics is described in Section 11.5 of *Fundamentals of Circuits and Filters*. The reverse transit time τ_R is very important to model the switching property of the lumped bipolar transistor, and it is a strong function of the biasing condition and temperature. Both phenomena are not implemented in the SPICE model.

1.1.8 SiGe HBTs

The performance of the Si bipolar transistor can be greatly enhanced with proper engineering of the base bandgap profile using a narrower bandgap material, SiGe, an alloy of Si and Ge. Structure wise, a SiGe

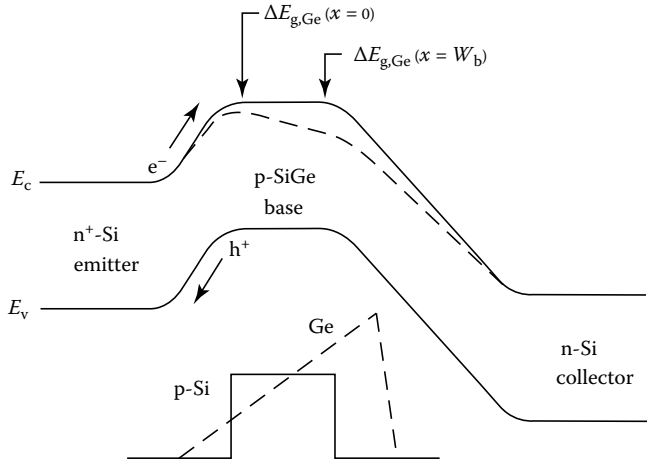


FIGURE 1.11 Energy band diagram of a graded base SiGe HBT and a comparably constructed Si BJT.

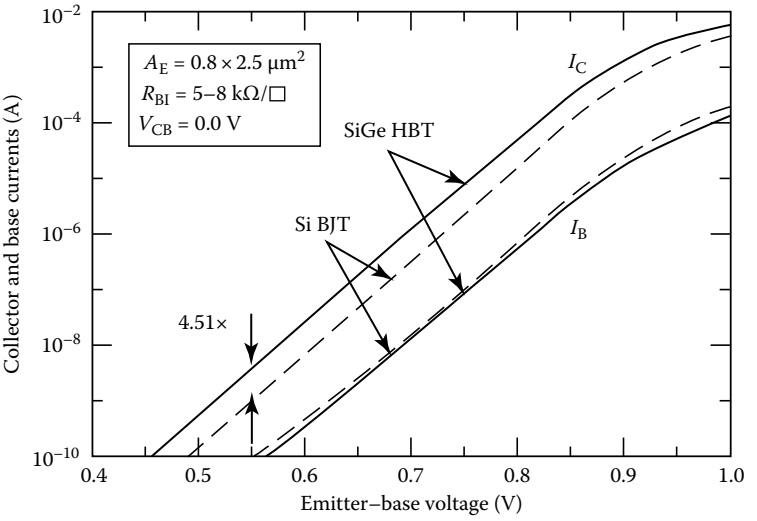


FIGURE 1.12 Experimental collector and base currents versus EB voltage for SiGe HBT and Si BJT.

HBT is essentially a Si BJT with a SiGe base. Its operation and circuit level performance advantages can be illustrated with the energy band diagram in Figure 1.11 [13]. Here the Ge content is linearly graded from emitter toward collector to create a large accelerating electric field that speeds up minority carrier transport across the base, thus making transistor speed much faster and cutoff frequency much higher. Everything else being the same, the potential barrier for electron injection into the base is reduced, thus exponentially enhancing the collector current. The base current is the same for SiGe HBT and Si BJT, as the emitter is typically made the same. Beta is thus higher in SiGe HBT. Figure 1.12 confirms these expectations experimentally with data from a typical first-generation SiGe HBT technology. The measured doping and Ge profiles are shown in Figure 1.13. The metallurgical base width is only 90 nm, and the neutral base width is around 50 nm. Figure 1.14 shows experimental cutoff frequency f_T improvement from using a graded SiGe base, which also directly translates into maximum oscillation frequency f_{max} improvement.

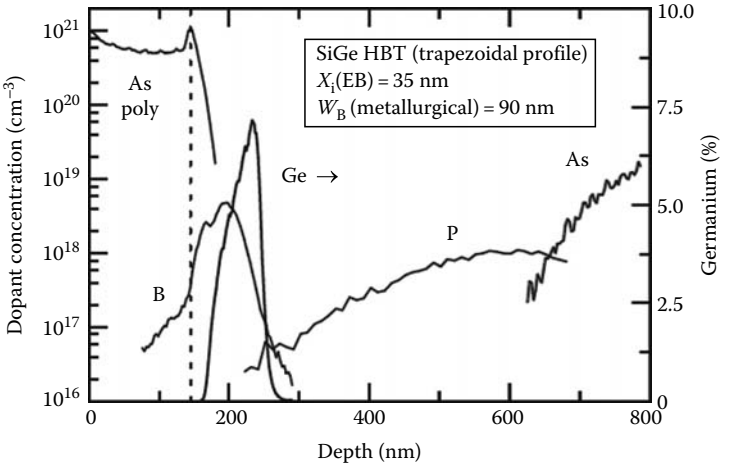


FIGURE 1.13 Measured doping and Ge profiles of a modern SiGe HBT.

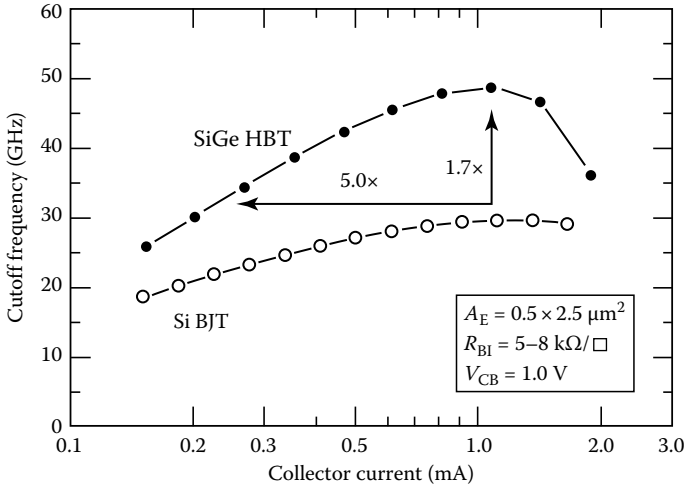


FIGURE 1.14 Experimental cutoff frequency versus collector current for SiGe HBT and Si BJT.

1.1.8.1 Operation Principle and Performance Advantages over Si BJT

In modern transistors, particularly with the use of polysilicon emitter, beta may be sufficient. If so, the higher beta potential of SiGe HBT can then be traded for reduced base resistance, through the use of higher base doping. The unique ability of simultaneously achieving high beta, low base resistance, and high cutoff frequency makes SiGe HBT attractive for many RF circuits. Broadband noise is naturally reduced, as low base resistance reduces transistor input noise voltage, and high beta as well as high f_T reduces transistor input noise current [13]. Experimentally, $1/f$ noise at the same base current was found to be approximately the same for SiGe HBT and Si BJT [14]. Consequently, $1/f$ noise is often naturally reduced in SiGe HBT circuits for the same biasing collector current, as base current is often smaller due to higher beta, as shown in Figure 1.15 using corner frequency as a figure-of-merit.

These, together with circuit-level optimization, can lead to excellent low-phase noise oscillators and frequency synthesizers suitable for both wireless and wire line communication circuits. Another less

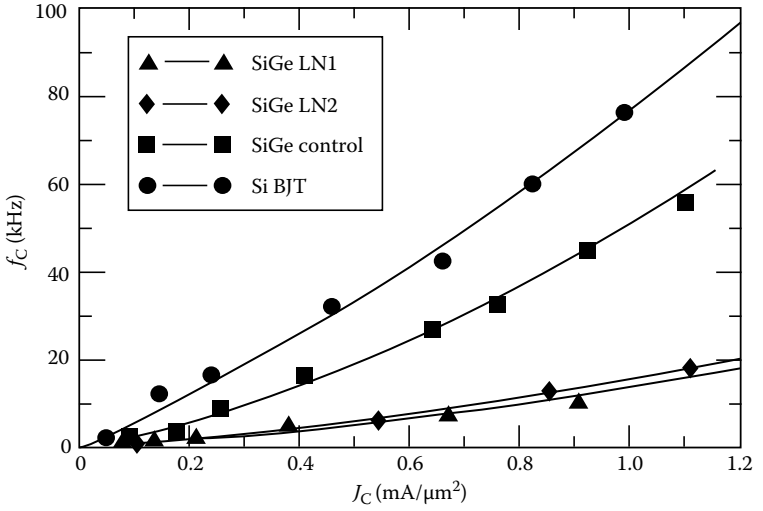


FIGURE 1.15 Experimentally measured corner frequency as a function of collector current density for three SiGe HBTs with different base SiGe designs, and a comparatively constructed Si BJT.

obvious advantage from grading Ge is the collector side of the neutral base has less impact on the collector current than the emitter side of the neutral base. Consequently, as collector voltage varies and the collector side of the neutral base is shifted toward the emitter due to increased CB junction depletion layer thickness, the collector current is increased to a much lesser extent than in a comparably constructed Si BJT, leading to a much higher output impedance or Early voltage. The $\beta \times V_A$ product is thus much higher in SiGe HBT than in Si BJT.

1.1.8.2 Industry Practice and Fabrication Technology

The standard industry practice today is to integrate SiGe HBT with CMOS, to form a SiGe BiCMOS technology. The ability to integrate with CMOS is also a significant advantage of SiGe HBT over III–V HBT. Modern SiGe BiCMOS combines the analog and RF performance advantages of the SiGe HBT, and the lower power logic, high integration level, and memory density of Si CMOS, into a single cost-effective system-on-chip (SoC) solution. Typically, SiGe HBTs with multiple breakdown voltages are offered through selective collector implantation, to provide more flexibility in circuit design.

The fabrication process of SiGe HBT and its integration with CMOS has been constantly evolving in the past two decades, and varies from company to company. Below are some common fabrication elements and modules shared by many if not all commercial first-generation (also most wide spread in manufacturing at present) SiGe technologies:

1. A starting N^+ subcollector around 5 Ω/sq on a p-type substrate at $5 \times 10^{15}/cm^3$, typically patterned to allow CMOS integration.
2. A high-temperature, lightly doped n-type collector, around 0.4–0.6 μm thick at $5 \times 10^{15}/cm^3$.
3. Polysilicon-filled deep trenches for isolation from adjacent devices, typically 1 μm wide and 7–10 μm deep.
4. Oxide filled shallow trenches or LOCOS for local device isolation, typically 0.3–0.6 μm deep.
5. An implanted collector reach through to the subcollector, typically at 10–20 Ωμm^2.
6. A composite SiGe epi layer consisting of a 10–20 nm Si buffer, a 70–100 nm boron-doped SiGe active layer, with or without C doping to help suppress boron out diffusion, and a 10–30 nm Si cap. The integrated boron dose is typically $1–3 \times 10^{13}/cm^2$.

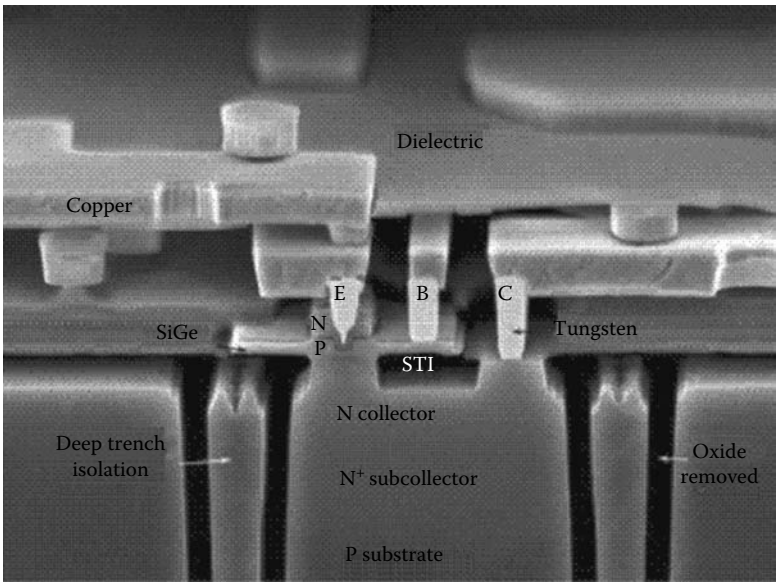


FIGURE 1.16 Structure of a modern SiGe HBT.

7. A variety of EB self-alignment scheme, depending on device structure and SiGe growth approach. All of them utilize some sort of spacer that is 100–300 nm wide.
8. Multiple self-aligned collector implantation to allow multiple breakdown voltages on the same chip.
9. Polysilicon extrinsic base, usually formed during SiGe growth over shallow trench oxide, and additional self-aligned extrinsic implantation to lower base resistance.
10. A silicided extrinsic base.
11. A 100–200 nm thick heavily doped ($>5 \times 10^{20}/\text{cm}^3$) polysilicon emitter, either implanted or in situ doped.
12. A variety of multiple level back-end-of-line metallization schemes using Al or Cu, typically borrowed from parent CMOS process.

These technological elements can also be seen in the electronic image of a second-generation SiGe HBT shown in Figure 1.16.

References

1. J. J. Ebers and J. M. Moll, Large signal behavior of bipolar transistors. *Proceedings IRE* 42, 1761–1772, December 1954.
2. H. K. Gummel and H. C. Poon, An integral charge-control model of bipolar transistors. *Bell System Technical Journal* 49, 827–852, May 1970.
3. L. W. Nagel and D. O. Pederson, SPICE (Simulation Program with Integrated Circuit Emphasis). University of California, Berkeley, ERL Memo No. ERL M382, April 1973.
4. P. Antognetti and G. Massobrio, *Semiconductor Device Modeling with SPICE*, McGraw-Hill, New York, 1988.
5. A. Vadimiresku, *The SPICE Book*, John Wiley & Sons, Hoboken, NJ, 1994.
6. A. S. Grove, *Physics and Technology of Semiconductor Devices*, John Wiley & Sons, Hoboken, NJ, 1967.
7. S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed., John Wiley & Sons, Hoboken, NJ, 1981.
8. G. W. Neudeck, *The PN Junction Diode*, Vol II, Modular Series on Solid-State Devices, Addison-Wesley, Upper Saddle River, NJ, 1983.

9. R. S. Muller and T. I. Kamins, *Device Electronics for Integrated Circuits*, 2nd ed., John Wiley & Sons, Hoboken, NJ, 1986.
10. E. S. Yang, *Microelectronic Devices*, McGraw-Hill, New York, 1988.
11. B. G. Streetman, *Solid State Electronic Devices*. 3rd ed., Prentice Hall, Upper Saddle River, NJ, 1990.
12. D. A. Neamen, *Semiconductor Physics and Devices*, Irwin, 1992.
13. J. D. Cressler and G. Niu, *Silicon–Germanium Heterojunction Bipolar Transistor*, Artech House, Norwood, MA, 2003.
14. G. Niu, Noise in SiGe HBT RF technology: Physics, modeling and circuit implications, *Proceedings of the IEEE*, pp. 1583–1597, September 2005.

1.2 Metal–Oxide–Silicon Field Effect Transistor

John Choma, Jr.

1.2.1 Introduction

Integrated electronic circuits realized in metal–oxide–silicon field effect transistor (MOSFET) technology are ubiquitous in both the commercial and military sectors of the technical community. To be sure, transistors manufactured in certain bipolar and III–V compound transistor technologies compete successfully with their MOSFET counterparts from such performance perspectives as switching speed, wideband frequency response, and insensitivity to electromagnetic interference and irradiated environments. Nevertheless, the MOSFET reigns supreme in the extant state of the electronics art for several reasons. The first of these reasons derives from the fact that the cross-section geometry of a MOSFET, when compared to that of most other solid-state transistors, is simpler. This simplicity affords a relative ease of foundry processing, which in turn promotes high device yield and therefore, cost-effective manufacturing. A second reason is that the surface area consumed on chip, or *footprint*, of a MOSFET is generally smaller than that of a comparably performing bipolar or III–V compound transistors. This feature allows increased packing density, which is particularly advantageous for digital signal processors that commonly require upwards of millions of transistors for system functionality. Third, MOSFETs can deliver acceptable circuit performance at low standby power levels, which is a laudable attribute in light of the aforementioned high device density digital architectures and the portability culture in which society is immersed presently. Finally, the native insulating oxide indigenous to the monolithic processing of silicon semiconductors renders MOSFET technologies amenable to the implementation of complex electronic systems on a single chip. No such native oxide prevails in III–V compound technologies, thereby rendering awkward the electrical isolation among the various components, subsystems, and subcircuits that comprise the overall electronic system.

The penchant toward adopting MOSFET technology for analog signal processing applications can also be rationalized. In particular, the nature of modern integrated systems is rarely exclusively digital or exclusively analog. Such systems are, in fact, ''mixed signal'' architectures that embody both digital and analog signal processing on the same chip. Because of the simplicity, packing density, and power dissipation attributes of MOSFETs, virtually 100% of digital architectures are realized in MOSFET technology. Prudence alone accordingly dictates a MOSFET technology realization of the analog cells implicit to a mixed signal framework if only to facilitate the electrical interface between the analog and digital units.

Aside from the operating flexibility and programmability advantages boasted by digital circuit schema, digital circuits in mixed signal architectures are often required to assure and sustain performance optimality of the analog signal flow paths in an electronic system. Unlike most digital networks, high-performance analog circuits are sensitive to specific values, or at least specific ranges of values, of several of the key physical and electrical parameters that effectively define the electrical properties of MOSFETs. Unfortunately, attaining the requisite accuracy in the numerical delineation of these parameters becomes

progressively more daunting as the performance metrics imposed on an analog network become more challenging and as device geometries scale to meet omnipresent quests for wider signal processing passbands. In these high-performance systems, digital subsystems are often deployed to sense the observable performance metrics of an analog signal flow path, compare said metrics to their respective optimal design goals, and then appropriately adjust the relevant electrical parameters or signal excitations implicit to the signal path. In effect, the combined digital controller and analog network behave as a seamless adaptive system that automatically corrects for manufacturing vagaries, increased device operating temperatures, and certain environmental effects.

The most commonly utilized MOSFETs in modern electronic systems come in two flavors: the N-channel MOSFET (NMOS), diagrammed in Figure 1.17 and the P-channel MOSFET (PMOS) shown in Figure 1.18. In the NMOS device of Figure 1.17, the bulk substrate is P-type and is doped to an average acceptor impurity concentration of N_A, for which a representative range of values is $5(10^{14})$ atoms/cm^3 $< N_A <$ 10^{16} atoms/cm^3. Its vertical depth, which is not expressly highlighted in the figure, is many times larger than the depth, Y_d, (of the order of a few tenths of microns) of either the source or drain diffusions or implants. These regions, whose widths are indicated as L_{diff} and which are connected electrically to the source (S) and drain (D) terminals of the MOSFET, are very strongly doped in that their donor impurity concentrations are $N_D = 10^{20}$ atoms/cm^3 or larger. The width, L_{diff}, is typically two- or three-times the channel length, indicated as L in the diagram. The metallization contact that forms the electrical terminal of the semiconductor bulk (B) is generally connected to the most negative potential available in the circuit into which the subject transistor is embedded. Such a connection reverse biases the PN junctions formed between the bulk and source regions and between the bulk and drain regions. This reverse biasing ensures that for at least low signal frequencies, the source and drain regions are electrically isolated from each other and from the bulk substrate. In certain types of multiwell IC processes, bulk–source and bulk–drain reverse biasing is assured simply by returning the bulk terminal directly to the source region contact.

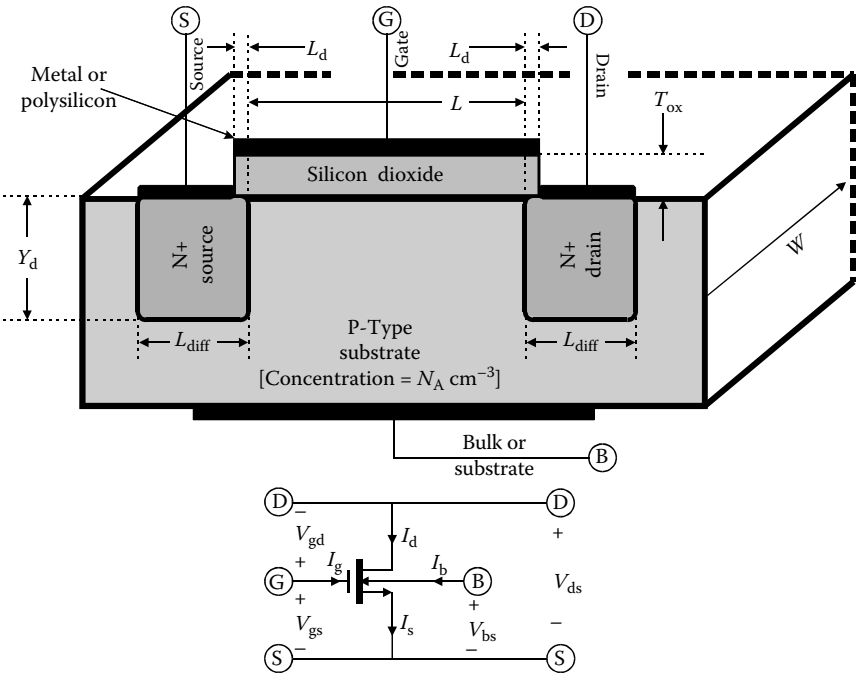


FIGURE 1.17 A simplified three-dimensional depiction of an N-channel MOSFET (NMOS) and its corresponding electrical schematic symbol. The diagram is not drawn to scale.

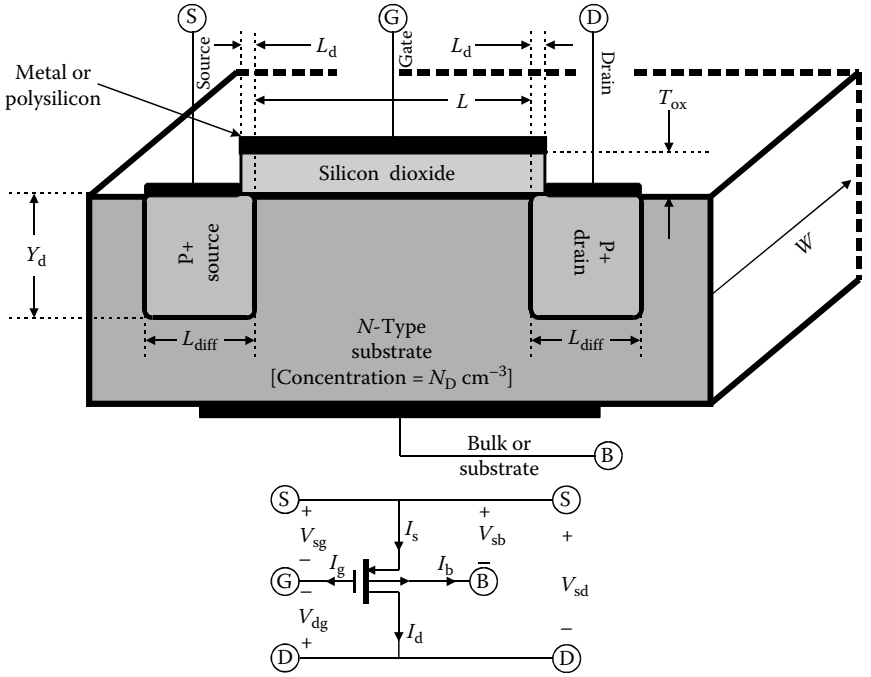


FIGURE 1.18 A simplified three-dimensional depiction of an P-channel MOSFET (PMOS) and its corresponding electrical schematic symbol. The diagram is not drawn to scale.

Lying atop the P-type bulk substrate is an insulating silicon dioxide layer of thickness T_{ox} that extends into the page as shown by a gate width, W. The oxide thickness in the extant state of the art is of the order of several tens of angstroms, where 1 Å is 10^{-8} cm. This oxide layer entirely covers the channel length, L, that separates the source region from the drain region, and it may overlap the source and drain regions by the amount, L_d, indicated in the diagram. The overlap of the source and drain regions is undesirable in that it limits broadband frequency responses in certain types of MOSFET amplifiers. In processes boasting self-aligned gate capabilities, L_d is ideally reduced to zero. But for state of the art processes delivering channel lengths as small as 65–130 nM, gate self-alignment focused on reducing L_d to no more than 5% of L is a challenging undertaking. The gate width, W, can be no smaller than the minimum channel length that can be produced by the identified foundry process. Subject to this proviso, the gate aspect ratio, W/L, is a designable parameter selected in accordance with the operating requirements of the circuit application for which the considered MOSFET is utilized.

The gate terminal (G) is formed by a contact made of a metallic or a polycrystalline silicon layer deposited directly atop the gate oxide. The gate metal of choice is aluminum. If the MOSFET under consideration is used in high-temperature environments and/or in applications that exploit low power supply voltages, polycrystalline silicon, which is commonly referred to as polysilicon, supplants the aluminum gate.

In addition to the simplified cross-section diagram of the N-channel MOSFET, Figure 1.17 inserts the electrical schematic symbol of the NMOS transistor. Of particular interest are the positive reference conventions adopted for four device currents and four device voltages. Specifically, positive drain current, I_d, flows into the transistor, as do the gate current, I_g, and the bulk, or substrate, current, I_b, while positive source current, I_s, flows out of the transistor. It follows from Kirchhoff's current law that

$$I_s = I_d + I_g + I_b. \tag{1.48}$$

However, since the gate contact is isolated from the semiconductor bulk by an insulating oxide layer, I_g is zero at the low frequencies for which capacitive phenomena associated with the insulating gate dielectric are insignificant. Moreover, the bulk current, I_b, is likewise almost zero at low signal frequencies, provided, as is usually the case, that care is taken to ensure reverse biasing of the bulk–drain and bulk–source PN junctions. Accordingly, the source and drain currents, I_s and I_d, respectively, are essentially identical when the frequencies of signals applied to the MOSFET are low. The pages that follow demonstrate that the static and low-frequency value of the drain, and hence the source, current is controlled by the gate-to-source voltage, V_{gs}, the drain-to-source voltage, V_{ds}, and, to a somewhat lesser extent, the bulk-to-source voltage V_{bs}. Stipulating an additional dependence of drain current on gate-to-drain voltage V_{gd} is superfluous, for by Kirchhoff's voltage law,

$$V_{ds} = V_{gs} - V_{gd}. \tag{1.49}$$

The P-channel MOSFET abstracted in Figure 1.18 is architecturally identical to its N-channel counterpart. The notable differences are that the bulk substrate in PMOS is N-type and the source and drain regions are heavily doped with P-type impurities. It follows that electrical isolation between the source region and the bulk, as well as between the drain region and the bulk, requires that the bulk substrate terminal of a PMOS device be connected either to the most positive of available circuit potentials or, if the process allows, to the source terminal. All of the geometrical parameters and their representative values remain the same as stipulated in conjunction with the NMOS unit. The PMOS electrical schematic symbol, which is also shown in the figure at hand, differs from the NMOS symbol in that the directions of the source terminal and bulk terminal arrows are reversed, as are the positive reference directions of all four transistor currents. While Equation 1.48 remains applicable, the analytical expression for the drain current, I_d, which now flows out of the transistor, is more conveniently couched in terms of the source-to-gate voltage, V_{sg}, the source-to-drain voltage, V_{sd}, and the source-to-bulk voltage, V_{sb}. The drain-to-gate voltage, V_{dg}, derives from

$$V_{sd} = V_{sg} - V_{dg}, \tag{1.50}$$

which mirrors Equation 1.49 subsequent to multiplying both sides of this equation by -1.

1.2.2 Channel Charge

A fundamental understanding of the physical charge storage and charge transport mechanisms that underpin the observable volt–ampere characteristics of considered transistors facilitates the reliable and reproducible design of high-performance analog networks in MOSFET technology. Aside from establishing a foundation upon which the static characteristic curves of a MOSFET can be constructed in a physically sound framework, these charge profiles also serve to define the voltage-dependent nature of the capacitance characteristics of a MOSFET. In effect, the subject charge profiles posture the MOSFET as a plausible varactor, which is useful in the monolithic design of voltage controlled oscillators, active filters, and other electronic networks.

The profile of charge stored in the channel between the source and drain regions of a MOSFET is best examined in terms of the simple circuit given in Figure 1.19a. In this circuit, the drain terminal is short circuited to the source to pin the drain–source voltage, V_{ds}, to zero. A zero bias is applied as indicated between the bulk and source, thereby establishing a charge depletion region about the PN junction formed between the substrate and source regions. Since the source and the drain are electrically connected to one another, the zero bias applied between bulk and source establishes an identical depletion zone about the bulk–drain PN junction. These depletion layers are delineated in the companion cross-section diagram of Figure 1.19b, as are the surface potential, φ_o, and the potential, V_{ox}, dropped across the gate silicon dioxide layer. With $V_{ds} = 0$, Equation 1.49 ensures a gate–source voltage, V_{gs}, that

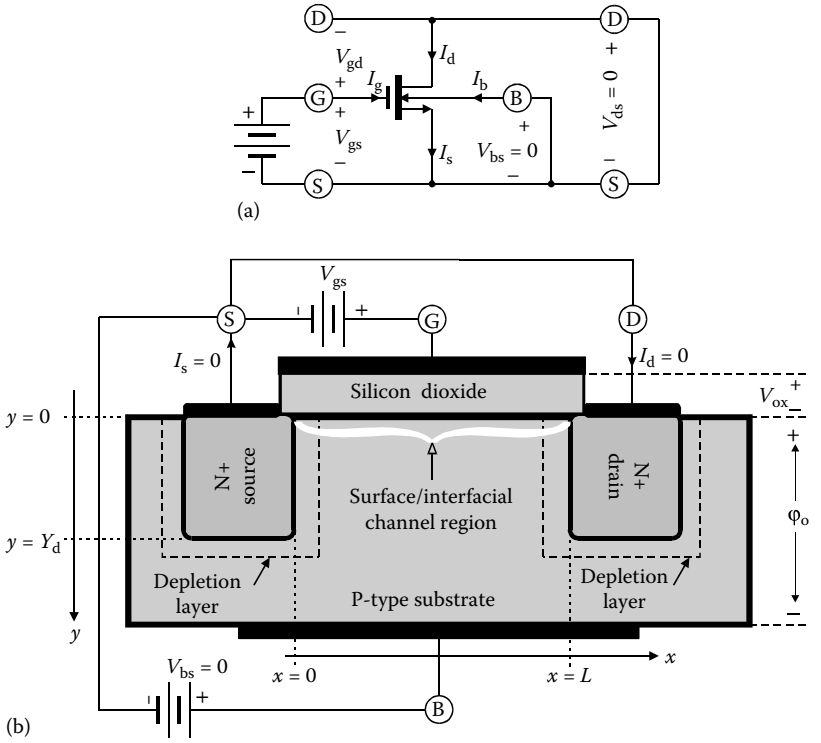


FIGURE 1.19 (a) NMOS transistor operated with $V_{ds} = 0$ and $V_{bs} = 0$. Although the battery connected between the gate and the source ensures $V_{gs} > 0$, $V_{gs} \leq 0$ is allowed in the discussion that references this circuit. (b) Cross-section diagram corresponding to the circuit in (a). Note that all applied voltages are referred to the source terminal. The diagram in (b) is not drawn to scale.

mirrors the gate–drain voltage, V_{gd}, regardless of the voltage applied between gate and source or gate and bulk terminals. In the absence of drain, source, bulk, and gate currents, $V_{ds} = 0$ also guarantees that surface potential φ_o, measured from the oxide semiconductor interface-to-the neutral zone of the bulk substrate, is the same throughout the channel region extending from $x = 0$-to-$x = L$ in the subject diagram. The aforementioned voltage, V_{ox}, includes the effects of parasitic trapped charge in the gate oxide, but it does not include the ramifications of work function differences that unavoidably prevail between the gate contact and the oxide and at the oxide–semiconductor interface. Note then that the voltage, V_{gb}, measured at the gate terminal with respect to the bulk terminal is, ignoring work function phenomena, simply

$$V_{gb} = V_{ox} + \varphi_o. \tag{1.51}$$

1.2.2.1 Surface Charge Density

A pivotally important analytical tool serving to define the charge, capacitance, and static volt–ampere characteristics of a MOSFET, is the charge density, $Q_o(\varphi_o)$, in units of coulombs per unit area, established at the semiconductor surface as a function of the surface potential, φ_o. Several authors have identified this charge profile as [1–3]

$$Q_o(\varphi_o) = -\text{sgn}(\varphi_o)\left(\frac{\sqrt{2}\varepsilon_s V_T}{D_b}\right)\sqrt{G(-\varphi_o) + G(\varphi_o)e^{-2V_F/V_T}}, \tag{1.52}$$

where $\varepsilon_s = 1.037$ pF/cm denotes the dielectric constant of silicon, and

$$\text{sgn}(\varphi_o) = \begin{cases} +1 & \text{for } \varphi_o > 0 \\ -1 & \text{for } \varphi_o < 0 \end{cases}. \tag{1.53}$$

In Equation 1.52,

$$V_T = kT/q \tag{1.54}$$

is the familiar semiconductor thermal voltage for which $k = (1.38)(10^{-23})$ J/K is Boltzmann's constant, $q = (1.60)(10^{-19})$ C is the magnitude of electron charge, and T is the absolute temperature of the semiconductor surface. The voltage, V_F, in the radical on the right-hand side of Equation 1.52 is the Fermi potential, which is given by

$$V_F \triangleq V_T \ln\left(\frac{N_A}{N_i}\right), \tag{1.55}$$

where N_A is the previously defined average acceptor impurity concentration of the bulk substrate in NMOS and $N_i = (1.45)(10^{10})$ atoms/cm^3 is the intrinsic carrier concentration of silicon at $T = 27°$C. The parameter, D_b, is known as the electron Debye length and is given by

$$D_b \triangleq \sqrt{\frac{\varepsilon_s V_T}{qN_A}}. \tag{1.56}$$

Finally, the function, $G(\varphi_o)$, in Equation 1.52 is

$$G(\varphi_o) = e^{\varphi_o/V_T} - 1 - \frac{\varphi_o}{V_T}, \tag{1.57}$$

where it is understood that the surface potential, φ_o, measured with respect to the charge neutral zone in the bulk in Figure 1.19 is established in response to an applied gate–bulk voltage, V_{gb}, or an applied gate–source voltage, V_{gs}. Observe that $G(\varphi_o) = G(-\varphi_o) = 0$ for $\varphi_o = 0$, which delivers the expected result in Equation 1.52 of $Q_o(0) = 0$. It should be understood that Equation 1.52 is premised on Poisson's equation and the Boltzmann carrier relationship,

$$p(0) = N_A e^{-\varphi_o/V_T}, \tag{1.58}$$

where $p(0)$ signifies the hole concentration at the surface if complete ionization of substrate dopant atoms is tacitly presumed. Since

$$p(0)n(0) = N_i^2, \tag{1.59}$$

the corresponding concentration of free surface electrons, $n(0)$, is

$$n(0) = \frac{N_i^2}{N_A} e^{\varphi_o/V_T} = N_A e^{(\varphi_o - 2V_F)/V_T}, \tag{1.60}$$

where Equation 1.55 has been exploited.

Because $G(\varphi_o)$ in Equation 1.57, as well as its companion relationship, $G(-\varphi_o)$, is a nonnegative number for all positive and negative values of the surface potential, the radical on the right-hand side of Equation 1.52 is a positive real number. Accordingly, Equation 1.53 forces $Q_o(\varphi_o) > 0$ for $\varphi_o < 0$ and $Q_o(\varphi_o) < 0$ for $\varphi_o > 0$. The positive nature of the surface charge density for negative surface potentials is indicative of bulk substrate holes attracted to the semiconductor surface because of the force exerted by the surface electric field established in response to negative surface potential. From Gauss' law, this field, say $E_o(\varphi_o)$ is simply

$$E_o(\varphi_o) = -\frac{Q_o(\varphi_o)}{\varepsilon_s} = \text{sgn}(\varphi_o)\left(\frac{\sqrt{2}V_T}{D_b}\right)\sqrt{G(-\varphi_o) + G(\varphi_o)e^{-2V_F/V_T}}, \tag{1.61}$$

which is indeed negative for $\varphi_o < 0$. Observe that Equation 1.58 supports the contention of an enhanced surface hole concentration when the potential established at the semiconductor surface is negative.

An equilibrium condition, which is more commonly referred to in the literature as the *flatband operating condition*, is reached when the applied gate–bulk or gate–source voltage produces a null surface potential, that is, $\varphi_o = 0$. For $\varphi_o = 0$, the net surface charge, $Q_o(\varphi_o)$, in Equation 1.52 is zero, as is the surface electric field, $E_o(\varphi_o)$, in Equation 1.61. Note further that by Equation 1.58, $p(0) = N_A$, which is the equilibrium hole concentration indicative of the NMOS bulk substrate for the transistor abstracted in Figure 1.17, assuming complete ionization of all substrate acceptor impurity atoms.

The negative surface charge prevailing for positive surface potentials, which gives rise to positive surface fields (field lines directed from the surface-to-the bulk substrate), reflects the surface charge depletion forged in response to holes repelled from the surface by $\varphi_o > 0$. Once again, Equation 1.58 is supportive of the proffered rationale in that it confirms a diminished surface hole concentration for progressively larger φ_o. Since departed holes leave in their wake a depletion zone of negative acceptor ions, the negative surface charge density resulting from positive surface potential is hardly surprising.

In addition to repelling holes from the semiconductor surface, Equation 1.60 indicates that the surface electron concentration increases as the surface potential, φ_o, rises above zero. Moreover, Equation 1.52 lends credence to this enhanced electron concentration claim since $Q_o(\varphi_o)$ is seen as becoming monotonically more negative as surface potential φ_o rises above zero. Indeed, the impact of the positive electric field associated with $\varphi_o > 0$ is to establish a force serving to attract the minority carriers (electrons) in the bulk substrate to the surface. For a surface potential in the range, $0 < \varphi_o < V_F$, the depletion charge contribution to the net surface charge continues to dominate over the charge associated with electrons cajoled to the surface, and the surface is said to operate in *depletion mode*. But as φ_o approaches and ultimately surpasses the Fermi potential, V_F, the impact on the nature of the surface charge becomes increasingly more interesting. For example, consider $\varphi_o = V_F$, for which Equations 1.58 and 1.60 yield $p(0) \equiv n(0) = N_i$, that is, the hole and electron concentrations at the surface are identically equal to the intrinsic carrier concentration. In effect, the surface region of the semiconductor changes from obviously P-type-to-intrinsic type, which is to say that the surface at $\varphi_o = V_F$ is neither P-type nor N-type.

For $V_F < \varphi_o < 2V_F$, Equations 1.58 and 1.60 project a surface electron concentration that actually exceeds the surface hole concentration, despite the originally P-type character of the semiconductor surface. In this range of surface potentials, the depletion layer at the surface continues to expand into the substrate but because of the enhanced electron concentration, the surface is said to operate in a condition of *weak inversion*. Weak inversion is significant from an engineering perspective in that it begins to establish the necessary condition for promoting observable drain and source current flow. In particular, suppose that the drain–source voltage, V_{ds}, were to be increased from its present null value to a suitably positive value. The presence of a significant mobile surface charge density in the form of free electrons allows said electrons to be transported from the source-to-the drain by the force associated with the lateral electric field established in response to $V_{ds} > 0$. In turn, this charge transport gives rise to a drain current flowing into the transistor and a source current flowing out of the device.

When φ_o rises to the value, $2V_F$, Equation 1.60 confirms a surface electron concentration that is numerically equal to the substrate doping concentration, N_A. In other words, the surface electron concentration precipitated by the strong positive electric fields implicit to $\varphi_o = 2V_F$ is identical to the equilibrium hole concentration evidenced in a silicon mass whose impurity concentration of completely ionized acceptor atoms is N_A. The surface has effectively changed its sex from its former P-type state to a field-induced (hence the terminology, ''field-effect,'' in the FET nomenclature) N-type state. Since the resultant surface electron concentration, $n(0)$, is rendered substantive, appreciable drain and source currents can flow for even modest values of applied drain–source voltages. In effect, the transistor can be said to be ''turned on'' when φ_o rises to twice the Fermi potential in the sense that a capability for substantial drain current flow is forged. When $\varphi_o \geq 2V_F$, the semiconductor surface is *strongly inverted*, or simply *inverted*.

Figure 1.20 displays a representative surface charge density profile as a function of the surface potential. Since a logarithmic charge scale is required to display all salient features of the charge density, the negative nature of the surface charge for positive surface potentials compels plotting the magnitude of the surface charge density on the vertical (charge) scale in the subject figure. The horizontal (voltage) scale is normalized to the thermal voltage, V_T. The plot invokes the presumptions of a 27°C semiconductor surface temperature and a substrate impurity concentration of $N_A = 10^{15}$ atoms/cm^3. For these stipulations, the thermal voltage is $V_T = 25.89$ mV, and the Fermi potential is $V_F = 288.4$ mV, whence $V_F/V_T = 11.14$. The plot displayed in Figure 1.20 clearly identifies the regions of hole accumulation ($\varphi_o < 0$), surface depletion ($0 < \varphi_o < 2V_F$), weak inversion, as typified by the increased concentration of free electrons at the surface ($V_F < \varphi_o < 2V_F$), and strong inversion, for which $\varphi_o \geq 2V_F$.

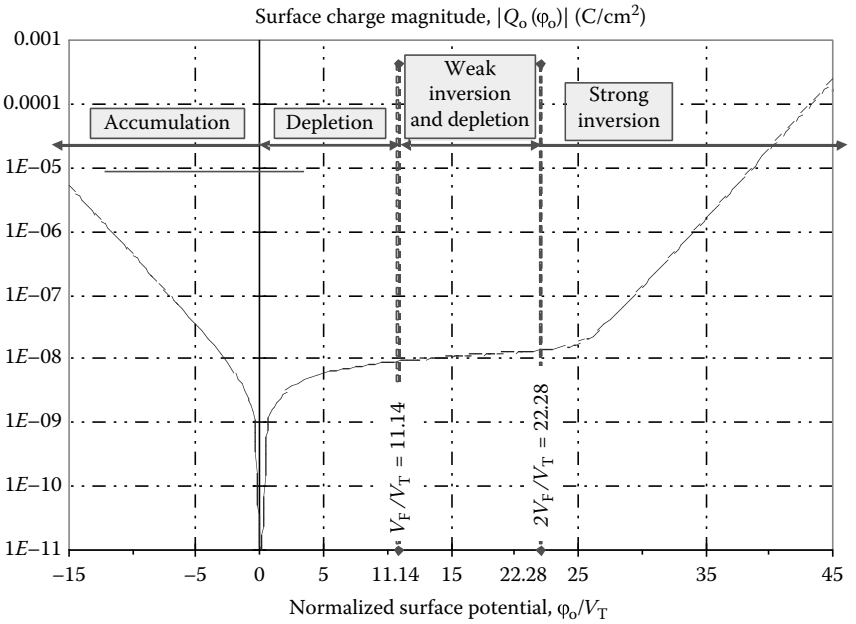


FIGURE 1.20 The magnitude of the surface charge density in the channel interfacial region for the MOSFET configured as shown in Figure 1.19b. A surface temperature of 27°C is assumed, as is a substrate impurity concentration of $N_A = 10^{15}$ atoms/cm^2.

1.2.2.2 Gate–Bulk Capacitance

The density of the net gate-to-bulk capacitance, $C_{gb}(\varphi_o)$, of the MOSFET whose cross-section diagram appears in Figure 1.19b is a series combination of the oxide capacitance density, C_{ox}, and the density of capacitance $C_d(\varphi_o)$, which is established between the oxide–substrate interface and the charge neutral region of the bulk. The pertinent equivalent circuit for $V_{ds} = 0$ is the structure depicted in Figure 1.21b, for which

$$C_{gb}(\varphi_o) = \frac{C_{ox}C_d(\varphi_o)}{C_{ox} + C_d(\varphi_o)}. \tag{1.62}$$

In Equation 1.62,

$$C_{ox} = \frac{\varepsilon_{ox}}{T_{ox}}, \tag{1.63}$$

where $\varepsilon_{ox} = 345$ fF/cm is the dielectric constant of silicon dioxide. Moreover,

$$C_d(\varphi_o) = \frac{d|Q_o(\varphi_o)|}{d\varphi_o}, \tag{1.64}$$

where the surface charge density, $Q_o(\varphi_o)$, is defined by Equation 1.52. After a trifle of differential calculus pain, it can be shown that

$$C_d(\varphi_o) = \text{sgn}(\varphi_o)\frac{\varepsilon_s}{\sqrt{2}D_b}\left[\frac{\left(e^{\varphi_o/V_T} - 1\right)e^{-2V_F/V_T} - \left(e^{-\varphi_o/V_T} - 1\right)}{\sqrt{G(-\varphi_o) + G(\varphi_o)e^{-2V_F/V_T}}}\right]. \tag{1.65}$$

The result at hand defines the surface capacitance density for all values of the surface potential, φ_o. A problem arises for $\varphi_o = 0$ in that the right-hand side becomes an indeterminate 0/0 form. This problem is circumvented by supplanting the exponential terms on the right-hand side of Equation 1.65, inclusive of those embedded in the functions, $G(\varphi_o)$ and $G(-\varphi_o)$, by their second order MacLaurin series expansions. Upon replacement of these exponential terms by said expansions, the surface capacitance density at the flatband condition, $\varphi_o = 0$, is found to be

$$C_d(0) = \frac{\varepsilon_s}{D_b}\sqrt{1 + e^{-2V_F/V_T}} \approx \frac{\varepsilon_s}{D_b} \triangleq C_{FB}, \tag{1.66}$$

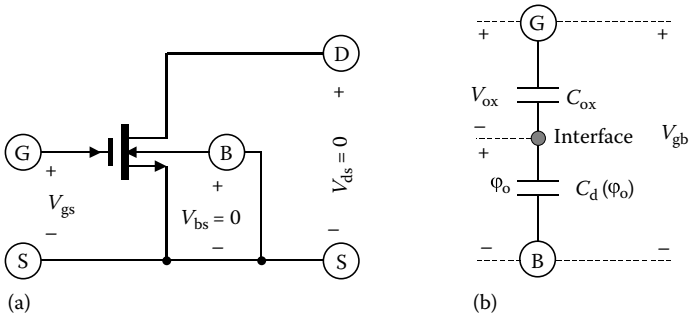


FIGURE 1.21 (a) NMOS transistor of Figure 1.19a operated with $V_{ds} = 0$ and $V_{bs} = 0$. (b) Circuit model between the gate and bulk terminals of the transistor in (a).

where C_{FB} is termed the surface flatband capacitance. The indicated approximation exploits the presumption that the impurity concentration, N_A, in the bulk substrate is significantly larger than the intrinsic carrier concentration, N_i, of silicon. It follows from Equation 1.62 that

$$\frac{C_{gb}(\varphi_o)}{C_{ox}} = \frac{1}{1 + \frac{C_{ox}}{C_d(\varphi_o)}}, \tag{1.67}$$

for which

$$\frac{C_{gb}(0)}{C_{ox}} \approx \frac{1}{1 + \left(\frac{\varepsilon_{ox}}{\varepsilon_s}\right)\left(\frac{D_b}{T_{ox}}\right)}. \tag{1.68}$$

Figure 1.22 displays a plot of the normalized gate–bulk capacitance, $C_{gb}(\varphi_o)/C_{ox}$, as a function of the normalized surface potential, φ_o/V_T, at room temperature (27°C) conditions. The relevant MOSFET is presumed to have an acceptor impurity concentration, N_A, in the bulk of 10^{15} atoms/cm^2, and a gate silicon dioxide thickness, T_{ox}, of 30 Å. The curve shows that in strong accumulation where $\varphi_o \ll 0$, the gate–bulk capacitance per unit area approaches the density of the oxide capacitance, C_{ox}. This observation reflects engineering expectations in that $\varphi_o \ll 0$ attracts a very large concentration of holes to the surface, for which the associated charge density serves to increase dramatically the surface density of capacitance, $C_d(\varphi_o)$. Indeed, for $\varphi_o \ll 0$, it is a simple matter to show that Equation 1.65 collapses to

$$C_d(\varphi_o)\big|_{\varphi_o \ll 0} \approx \frac{\varepsilon_s}{\sqrt{2}D_b}\sqrt{e^{-\varphi_o/V_T} - 1} \approx \frac{\varepsilon_s}{\sqrt{2}D_b}e^{|\varphi_o|/2V_T}, \tag{1.69}$$

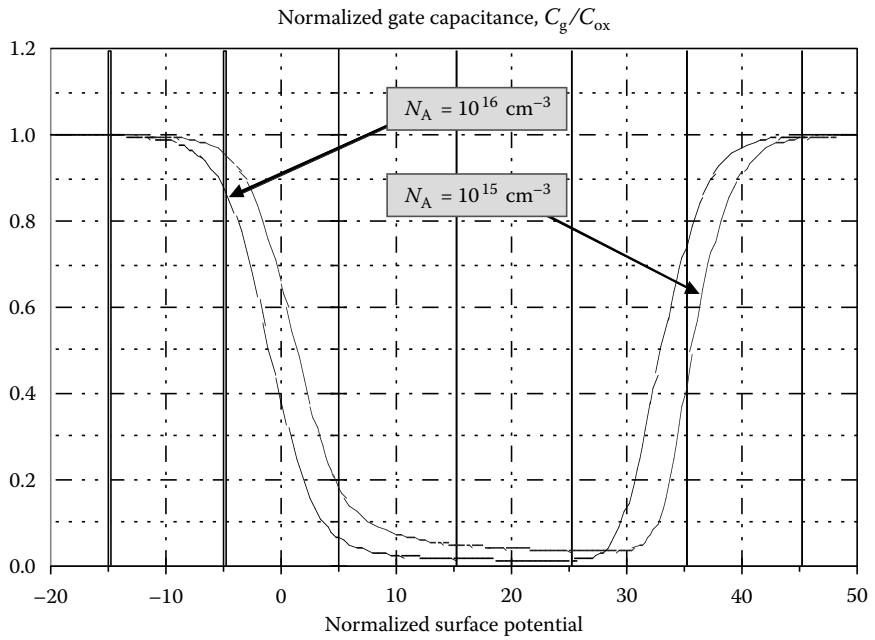


FIGURE 1.22 The normalized gate-to-bulk capacitance of the N-channel MOSFET shown in Figure 1.21a as a function of the indicated normalized surface potential. The temperature of the oxide–semiconductor interface is taken to be $T = 27$°C, the oxide thickness is $T_{ox} = 30$ Å, and the acceptor impurity concentration in the bulk is $N_A = 10^{15}$ atoms/cm^2. The curve is applicable to only low signal frequencies.

which clearly suggests a sharp rise in capacitance density with the absolute value of the negative surface potential. Since the surface capacitance density can be viewed as a ratio of the silicon dielectric constant, ε_s, to an effective and voltage-dependent dielectric thickness, say $y(\varphi_o)$, observe a dielectric thickness associated with Equation 1.69 of

$$y(\varphi_o)\big|_{\varphi_o \ll 0} \approx \sqrt{2} D_b e^{-|\varphi_o|/2V_T}, \tag{1.70}$$

which diminishes rapidly with progressively more negative surface potentials.

As the surface potential increases toward and beyond zero, the normalized capacitance plotted in Figure 1.22 decreases because a depletion layer begins to form at the interface. This depletion layer acts as a dielectric whose thickness increases as a nominal square root function of the surface potential. Under the depletion condition, the surface capacitance given by Equation 1.65 can be approximated by

$$C_d(\varphi_o)\big|_{\text{Depletion}} \approx \frac{\varepsilon_s}{\sqrt{2} D_b \sqrt{\frac{\varphi_o}{V_T} - 1}}, \tag{1.71}$$

which implies a depletion layer thickness, say y_d, (not to be confused with Y_d, the depth of the source and drain regions) of

$$y_d \triangleq y(\varphi_o)\big|_{\text{Depletion}} \approx D_b \sqrt{2\left(\frac{\varphi_o}{V_T} - 1\right)}. \tag{1.72}$$

Equation 1.71 approximates the actual surface capacitance density to within an error magnitude of nominally less than 10% for $3V_T \le \varphi_o \le 20V_T$. This allowable range of surface potential can actually be extended to embrace $\varphi_o \le 2V_F$ since for $\varphi_o > V_F$, an appreciable portion of the charge observed at the interface can be attributed to free electrons, and not simply to the ionic charge in the depletion layer forged by holes repelled from the interface.

As φ_o continues to increase, $C_d(\varphi_o)$, and hence $C_{gb}(\varphi_o)$, continues decreasing toward a minimum value that is achieved at a value close to a surface potential of $2V_F$, which is the threshold of strong surface inversion. At this potential, the thickness of the depletion layer implicit to the interfacial capacitance density is maximized since further increases in the surface charge density derive dominantly from electrons attracted to the surface. Rather than indulge in the academic propriety of using Equation 1.65 to compute the exact surface potential commensurate with minimal surface capacitance density, engineering prudence encourages the simplified approach of presuming $\varphi_o = 2V_F$ to be a sufficiently accurate requirement for minimal depletion capacitance. Upon adoption of this stance, Equation 1.72 is suitable for computing the maximum thickness, say W_d, of the depletion layer. Accordingly,

$$W_d \triangleq y(2V_F) \approx D_b \sqrt{2\left(\frac{2V_F}{V_T} - 1\right)}, \tag{1.73}$$

and since $2V_F$ is invariably much larger than the thermal voltage, V_T, Equation 1.56 allows this result to be written as

$$W_d \approx 2\sqrt{\frac{\varepsilon_s V_F}{qN_A}}. \tag{1.74}$$

The resultant minimum density of surface depletion capacitance is

$$C_{\rm d}(2V_{\rm F}) = \frac{\varepsilon_{\rm s}}{W_{\rm d}} \approx \frac{1}{2}\sqrt{\frac{qN_{\rm A}\varepsilon_{\rm s}}{V_{\rm F}}}. \tag{1.75}$$

Using Equations 1.67 and 1.63, the corresponding density, $C_{\rm min}$, of minimum gate–bulk capacitance is

$$C_{\rm min} \approx C_{\rm gb}(2V_{\rm F}) \approx \frac{\varepsilon_{\rm ox}/T_{\rm ox}}{1 + 2\left(\frac{\varepsilon_{\rm ox}}{T_{\rm ox}}\right)\sqrt{\frac{V_{\rm F}}{qN_{\rm A}\varepsilon_{\rm s}}}}. \tag{1.76}$$

Observe that the maximum factor by which the effective gate–bulk capacitance can be reduced is

$$\frac{C_{\rm ox}}{C_{\rm min}} \approx 1 + 2\left(\frac{\varepsilon_{\rm ox}}{T_{\rm ox}}\right)\sqrt{\frac{V_{\rm F}}{qN_{\rm A}\varepsilon_{\rm s}}}, \tag{1.77}$$

which involves parameters that are largely out of the control of the circuit designer. By inspection of Figure 1.22, this capacitance perturbation requires a surface potential swing extending from roughly $-15V_{\rm T}$ (about -400 mV at 27°C) to $2V_{\rm F}$ (generally smaller than 600 mV). Although the requisite surface potential excursion is somewhat large for maximal capacitance modulation, it should be noted that the maximum capacitance change factor predicted by Equation 1.77 can be as large as almost 100.

As $\varphi_{\rm o}$ increases beyond twice the Fermi potential, the interface charge density increases robustly as the semiconductor surface begins to invert strongly. Figure 1.22 resultantly displays an increasing bulk–gate capacitance density, not unlike the increased capacitance prevailing in strong accumulation because of holes attracted to the interface. Under actual measurement conditions, however, the indicated increased capacitance for $\varphi_{\rm o} > 2V_{\rm F}$ is observed only when the frequencies of signals established between the gate and bulk are below a few tens of hertz [4–6]. The problem is that for most practical signal frequencies, the recombination–generation rates of electrons in NMOS devices are unable to track with the signal-induced exchanges in charge between the neutral bulk and the inversion layer. Figure 1.23 displays the true gate–bulk capacitance characteristics for practical signal frequencies, wherein the dashed segment drawn for $\varphi_{\rm o} > 2V_{\rm F}$ is the applicable high-voltage capacitance trace for frequencies above a few tens of hertz.

1.2.2.3 Approximate Depletion Zone Analysis

As delineated in the discussion pertaining to the surface charge density defined by Equation 1.52, the MOSFET in Figure 1.21a exhibits depletion at the oxide–semiconductor interface for surface potentials satisfying the constraint, $0 < \varphi_{\rm o} < 2V_{\rm F}$. The formation of the surface depletion zone is critically important to the establishment of the volt–ampere characteristics of a MOSFET because it serves as a precursor to the surface inversion that comprises the necessary condition for drain and source current conduction. Recall, for example, that weak inversion is said to initiate at $\varphi_{\rm o} = V_{\rm F}$, in the sense that the original P-type character of the interfacial semiconductor is transformed to intrinsic material. When $\varphi_{\rm o}$ is elevated to $2V_{\rm F}$, the surface is strongly inverted in that the concentration of free electrons at the surface increases to a value that is identical to the average impurity concentration in the substrate. Although the concentration of surface electrons begins to increase for $\varphi_{\rm o}$ barely above zero, as is highlighted by Equation 1.60, an electron concentration commensurate with the possibility of substantial drain and source current flow does not materialize until the surface potential, $\varphi_{\rm o}$, reaches the immediate neighborhood of twice the Fermi potential. An attribute of Equation 1.52 is that this relationship does not explicitly distinguish between immobile depletion charge and mobile electron charge, both of which contribute to the observed surface charge density. A shortfall of Equation 1.52 is that its analytically cumbersome nature all but precludes the development of mathematically tractable expressions for the volt–ampere characteristics of

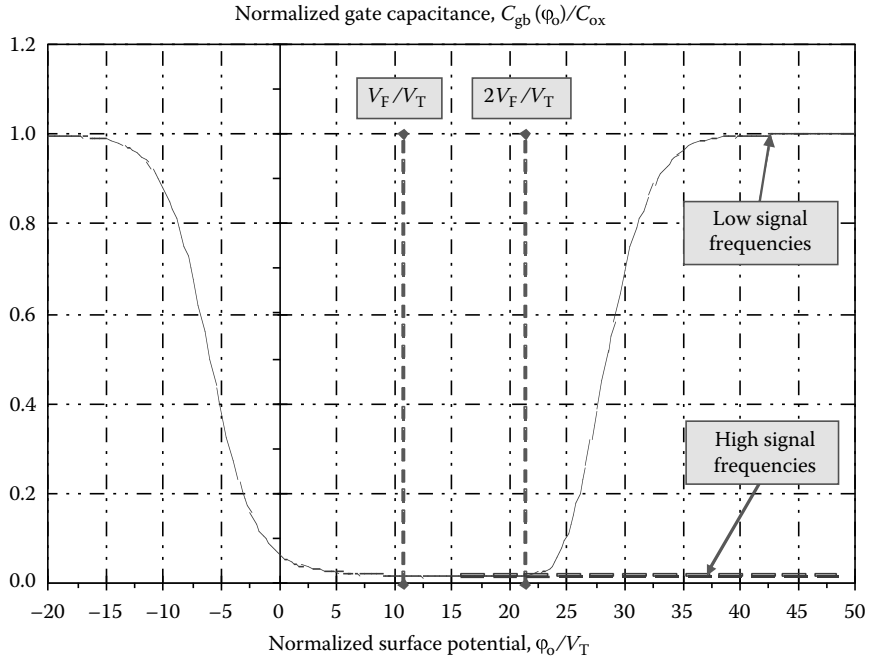


FIGURE 1.23 The capacitance characteristics of Figure 1.22 for the conditions of both low and high signal frequencies.

a MOSFET. Fortunately, the awkwardness of Equation 1.52 is mitigated if the reasonable approximation is made that for $0 < \varphi_o < 2V_F$, the charge in the surface channel region derives exclusively from depletion phenomena, that is, a substantive density of free electron charge does not materialize at the surface until $\varphi_o = 2V_F$.

To the extent that the entire substrate region is uniformly doped at the indicated impurity concentration of N_A and assuming complete ionization of all substrate impurity atoms, the resultant concentration, say $\rho(y)$, of immobile ionic charge in the depletion zone throughout the channel region from source-to-drain is nominally constant at the value, $-qN_A$. Of course, the electron concentration at the surface increases in proportion to the decreased hole population therein but as long as φ_o remains smaller than $2V_F$, Equation 1.60 confirms that the free electron concentration is significantly smaller than N_A.

Figure 1.24a depicts the depletion charge density, $\rho(y)$, beneath the oxide–semiconductor interface, where W_d represents the depth of the depletion layer established at the interface. Using Gauss' law, the electric field, $E(y)$, promoted by this charge concentration profile derives from

$$\frac{dE(y)}{dy} = \frac{\rho(y)}{\varepsilon_s}. \tag{1.78}$$

Since $\rho(y) = -qN_A$ for $0 \le y \le W_d$, Equation 1.78 implies

$$\int_{E(y)}^{E(W_d)} dE(y) = -\frac{qN_A}{\varepsilon_s} \int_{y}^{W_d} dy. \tag{1.79}$$

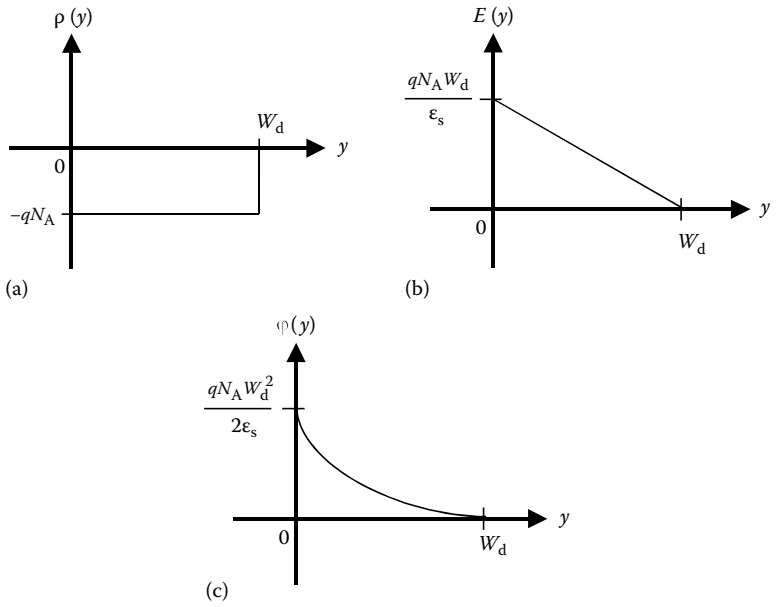


FIGURE 1.24 (a) The approximate profile of the depletion charge concentration at the surface of the MOSFET depicted in Figure 1.21a. (b) The electric field intensity as a function of bulk substrate depth measured with respect to the interfacial surface, corresponding to the charge profile in (a). (c) The potential implied by the electric field plot in (b).

In view of the fact that $E(W_d)$ is zero in the undepleted, charge neutral substrate region corresponding to $W \geq W_d$, Equation 1.79 produces the linear electric field relationship,

$$E(y) = \frac{qN_A W_d}{\varepsilon_s}\left(1 - \frac{y}{W_d}\right) = \frac{V_T W_d}{D_b^2}\left(1 - \frac{y}{W_d}\right), \tag{1.80}$$

where Equation 1.56 for the Debye length is invoked. Equation 1.80 is sketched as a function of the substrate depth variable, y, in Figure 1.24b. Observe that maximum field intensity prevails at surface where

$$E(0) \triangleq E_o = \frac{V_T}{D_b}\left(\frac{W_d}{D_b}\right). \tag{1.81}$$

The potential, $\varphi(y)$, corresponding to the field intensity, $E(y)$, stipulated by Equation 1.80 satisfies

$$-\frac{d\varphi(y)}{dy} = E(y) = \frac{qN_A}{\varepsilon_s}(W_d - y). \tag{1.82}$$

If zero reference potential is ascribed to the substrate depth, W_d, beyond which the substrate is charge neutral, Equation 1.82 sets forth

$$\int_{\varphi(y)}^{0} d\varphi(y) = \frac{qN_A}{\varepsilon_s}\int_{y}^{W_d}(W_d - y)dy, \tag{1.83}$$

whence the bulk substrate potential, $\varphi(y)$, referenced to the potential evidenced at the substrate depletion depth, W_d, is

$$\varphi(y) = \frac{qN_A W_d^2}{2\varepsilon_s}\left(1 - \frac{y}{W_d}\right)^2 = \frac{1}{2}V_T\left(\frac{W_d}{D_b}\right)^2\left(1 - \frac{y}{W_d}\right)^2, \tag{1.84}$$

whose functional dependence on variable y is sketched in Figure 1.24c. Equation 1.84 suggests that the surface potential, $\varphi(0)$, which is effectively the net voltage dropped across the depleted region of the bulk substrate, is

$$\varphi(0) \triangleq \varphi_o = \frac{qN_A W_d^2}{2\varepsilon_s} = \frac{1}{2}V_T\left(\frac{W_d}{D_b}\right)^2 \equiv \frac{E_o W_d}{2}, \tag{1.85}$$

Since the interface potential, φ_o, and hence the depletion depth, W_d, is controlled externally by the applied gate-to-bulk voltage, V_{gb}, it is of interest to determine φ_o as an explicit function of V_{gb}. To this end, the electric field intensity, E_{ox}, in the silicon dioxide layer of the structure of Figure 1.19b is uniform throughout the oxide thickness by virtue of the insulating nature of the oxide. Ignoring work function phenomena prevailing between the gate contact and gate oxide, as well as between the oxide and semiconductor surface, this field is simply

$$E_{ox} = \frac{V_{gb} - \varphi_o}{T_{ox}}. \tag{1.86}$$

Equation 1.86 also invokes the approximation that the voltage dropped from the bottom of the depletion region-to-the bulk terminal is essentially zero. This assumption is reasonable in that the substrate is ultimately reverse biased to preclude substantive bulk current flow. Moreover, the holes displaced from the interface region-to-the neutral bulk render the neutral substrate zone a low resistivity volume. Because the field immediately below the interface is E_o, as defined by Equation 1.81, continuity constraints mandate

$$\varepsilon_{ox}E_{ox} = \varepsilon_s E_o, \tag{1.87}$$

or

$$\varepsilon_{ox}\left(\frac{V_{gb} - \varphi_o}{T_{ox}}\right) = \varepsilon_s\left(\frac{V_T W_d}{D_b^2}\right) \approx C_{FB}\left(\frac{W_d}{D_b}\right)V_T. \tag{1.88}$$

Armed with Equations 1.86, 1.63, and 1.85, parameter W_d in Equation 1.88 can be eliminated to arrive at the utilitarian expression,

$$V_{gb} = \varphi_o + \sqrt{2V_\theta\varphi_o}, \tag{1.89}$$

where the voltage metric, V_θ, termed the *body effect voltage*, is given by

$$V_\theta = V_T\left(\frac{C_{FB}}{C_{ox}}\right)^2 = V_T\left(\frac{\varepsilon_s}{\varepsilon_{ox}}\right)^2\left(\frac{T_{ox}}{D_b}\right)^2 = \frac{qN_A\varepsilon_s}{C_{ox}^2}. \tag{1.90}$$

Parameter V_θ is generally of the order of the mid-tens of microvolts.* Observe that V_θ is proportional to the square of the oxide thickness, T_{ox} and is therefore reduced sharply with diminishing gate oxide thickness.

The final step to the problem of determining the dependence of interface potential φ_o on applied gate-to-bulk voltage V_{gb} involves a straightforward solution of Equation 1.89 for φ_o. The result is

$$\varphi_o = V_{gb} + V_\theta - \sqrt{V_\theta\left(2V_{gb} + V_\theta\right)}. \tag{1.91}$$

As expected, $\varphi_o = 0$ for $V_{gb} = 0$. Ordinarily, V_θ is much smaller than practical values of the gate-to-bulk voltage, V_{gb}, so that Equation 1.91 can be approximated as

$$\varphi_o \approx V_{gb} - \sqrt{2V_\theta V_{gb}}, \tag{1.92}$$

which is similar in form to Equation 1.89.

1.2.2.4 Threshold

The approximate depletion regime analysis executed in Section 1.2.2.3 conveniently precipitates an analytical definition of the threshold condition, that is, the condition whereby strong inversion materializes at the oxide–semiconductor interface. It has been demonstrated that the onset of the threshold condition corresponds to a surface potential, φ_o, of twice the Fermi potential. Accordingly, threshold requires that the gate–bulk voltage V_{gb}, in Equation 1.89 rise to a value, say V_{gbh}, such that

$$V_{gbh} = 2V_F + 2\sqrt{V_\theta V_F}. \tag{1.93}$$

It is to be understood that a gate-to-bulk voltage, V_{gb}, satisfying the constraint, $V_{gb} \geq V_{gbh}$, is commensurate with instilling strong inversion at the surface of the MOSFET depicted in Figure 1.21a. On the presumption that the depth, W_d, of the depletion layer, corresponding to $\varphi_o = 2V_F$ is unaltered by gate–bulk voltage increases beyond the threshold value, V_{gbh}, the resultant charge profile offered in Figure 1.24a changes into the form diagrammed in Figure 1.25. In this diagram, N_s is the concentration of free

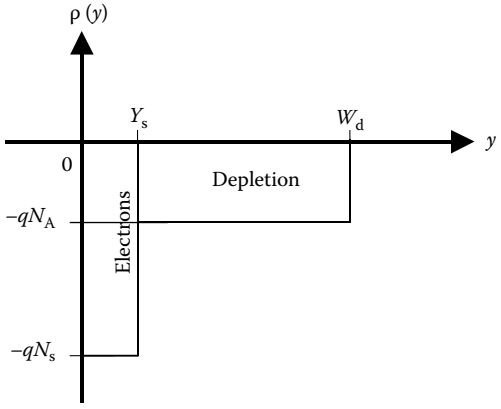


FIGURE 1.25 The approximate profile of the charge concentration for a strongly inverted surface in the MOSFET depicted in Figure 1.21a.

* Most HSPICE and other SPICE simulators use a *body effect parameter* to compute the extent to which the bulk–source voltage perturbs the gate–source threshold voltage. This body effect parameter, γ, derives from $\gamma^2 = 2V_\theta$.

electrons in the surface inversion layer, and Y_s is the thickness of the inversion layer. Depending on the value of the gate–bulk voltage excess, $(V_{gb} - V_{gbh})$, Y_s is typically 20%–50% larger than the electron Debye length.

Two circumstances limit the utility of Equation 1.93. The first of these is that MOSFETs are often operated with nonzero bulk–source bias, as opposed to the zero bias presumed to this juncture. If a bulk–source voltage, V_{bs}, is applied to the MOSFET in Figure 1.21a, the potentials at both the oxide–semiconductor interface and the bottom of the depletion layer are elevated by an amount, V_{bs}, which, in concert with earlier admonitions, is invariably a negative voltage to ensure reverse biasing of the bulk–source and bulk–drain PN junctions. Thus, the results documented in Section 1.2.2.3 remain valid because the voltage developed across the depletion layer is still the surface potential, φ_o, exploited therein. However, since the surface potential rises by V_{bs}, one of the necessary modifications to results disclosed earlier is that Equation 1.60 for the free electron concentration at the surface must be modified as

$$n(0) = \frac{N_i^2}{N_A} e^{(\varphi_o + V_{bs})/V_T} = N_A e^{(\varphi_o + V_{bs} - 2V_F)/V_T}. \tag{1.94}$$

Recall that the measure for the onset of strong inversion in a MOSFET is a surface electron concentration, $n(0)$, that equals to the hole concentration, N_A, in the equilibrium substrate. In order to effect this strong inversion condition, Equation 1.94 suggests the necessity of a surface potential that is at least as large as $(2V_F - V_{bs})$, as opposed to merely $2V_F$. Accordingly, the effect of bulk–source biasing on the gate–bulk threshold voltage can be embraced by replacing the voltage, $2V_F$, in Equation 1.93 by the voltage $(2V_F - V_{bs})$ so that

$$V_{gbh} = (2V_F - V_{bs}) + \sqrt{2V_\theta(2V_F - V_{bs})}. \tag{1.95}$$

The second shortfall of Equation 1.93 stems from the fact in actual circuit design environments, it is far more convenient to stipulate the minimum gate–source voltage, V_{gs}, and not the minimum gate–bulk voltage, V_{bs}, that establishes the onset of strong inversion. Since V_{gs} is the voltage sum, $(V_{gb} + V_{bs})$, adding V_{bs} to both sides of Equation 1.95 delivers a gate-to-source threshold voltage, say V_h, of the form,

$$V_h = V_{ho} + 2\sqrt{V_\theta V_F}\left(\sqrt{1 - \frac{V_{bs}}{2V_F}} - 1\right), \tag{1.96}$$

where V_{ho}, which represents the zero bias ($V_{bs} = 0$) value of the gate–source threshold potential, is

$$V_{ho} = 2\left(V_F + \sqrt{V_\theta V_F}\right). \tag{1.97}$$

In practice, the zero bias value, V_{ho}, of gate–source threshold voltage is best evaluated through measurement since it is strongly influenced by gate region work function phenomena and parasitic charges trapped in the gate oxide layer, whose engineering effects are difficult to quantify accurately and reliably. Observe in Equation 1.96 that the effect of an increasing bulk-to-source reverse bias ($V_{bs} < 0$) is to increase the threshold voltage above its zero bias value, V_{ho}, as a square root function of V_{bs}. This bulk-induced modulation of the threshold potential is rendered small by small values of the square root of parameter V_θ, which Equation 1.90 projects as directly dependent on the gate oxide thickness, T_{ox}. One reason for the current penchant toward progressively decreased oxide thickness is the minimization of threshold voltage modulation, which is generally an undesirable effect in MOSFETs deployed in analog network applications.

1.2.3 Volt–Ampere Characteristics

The static volt–ampere characteristics of the N-channel MOSFET in Figure 1.17 stipulate the dependence of the static drain current, I_d, on the static values of the gate–source voltage, V_{gs}, the drain–source voltage, V_{ds}, and the bulk–source voltage, V_{bs}, which is invariably a nonpositive voltage. It is convenient to partition these characteristics into three segments; namely, the *cutoff regime*, the *ohmic regime*, and the *saturation regime*.

1.2.3.1 Cutoff Regime

The cutoff regime is the most boring of the three MOSFET operating domains in that no drain current flows in cutoff, despite all reasonable positive value of the drain–source voltage, V_{ds}. Since drain current conduction requires surface inversion at the oxide–substrate interface, zero current is assured when no such charge inversion prevails. In turn, no inversion layer is formed when the gate–source voltage, V_{gs}, lies below its threshold value, V_h. Thus, in cutoff,

$$I_d = 0, \quad \text{if } V_{gs} < V_h. \tag{1.98}$$

As V_{gs} rises above zero but remains below V_h, a subthreshold current is induced from the drain-to-the source regions for $V_{ds} > 0$ [7]. This current is manifested by the fact that, as is conveyed by Equation 1.94, the interfacial free electron concentration increases slightly for surface potentials above 0 V. However, the current evidenced in the subthreshold regime is small because of the limited availability of free surface electrons. As a result, the gain and frequency response, in addition to the actual drain current, that the transistor is capable of mustering are limited. Although subthreshold operation enjoys utility in certain types of low-power system applications, such as hearing aids, where neither gain nor bandwidth are daunting requirements, it is rarely exploited in broadband and other high-performance applications.

1.2.3.2 Ohmic Regime

In the ohmic regime, which is sometimes called the *triode regime*, $V_{gs} \geq V_h$ and $V_{ds} \leq (V_{gs} - V_h)$. Thus, the interfacial surface of a MOSFET is strongly inverted in the ohmic domain, and simultaneously, a relatively small voltage is applied from the drain-to-the source. The voltage difference, $(V_{gs} - V_h)$, is commonly referenced as the *drain saturation voltage*, V_{dsat}, that is,

$$V_{dsat} \triangleq V_{gs} - V_h. \tag{1.99}$$

Because of Equation 1.49, observe that the provision, $V_{ds} \leq (V_{gs} - V_h)$, is equivalent to the requirement, $V_{gd} \geq V_h$. This is to say that a MOSFET operates in the ohmic regime if and only if both the gate–source and the gate–drain voltages are larger than the threshold potential. Viewed in yet another fashion, $V_{gs} \geq V_h$ and $V_{gd} \geq V_h$ ensure that both the source and the drain ends of the interfacial surface between the source and the drain regions are strongly inverted. A conduit, or channel, of free electrons that electrically couples the source to the drain is thereby established.

The aforementioned channel of free electrons is highlighted in the device cross section abstracted in Figure 1.26. Because $V_{gs} \geq V_h$ and $V_{gd} \geq V_h$, the electron inversion layer extends throughout the entire surface region from the source-to-the drain. But since $V_{gd} = (V_{gs} - V_{ds})$ and $V_{ds} > 0$, the gate-to-drain bias, V_{gd}, is necessarily smaller than its gate-to-source counterpart, V_{gs}. It follows that the surface potential in the neighborhood of the drain region is smaller than that prevailing near the source region, whence the electron concentration near the drain is smaller than it is at the source. Accordingly, the channel of electrons depicted in the figure at hand does not have a uniform depth (y-direction) and is, in fact, deeper at the source site, where $x = 0$, than it is at the drain site, which is typified analytically by $x = L$. For analogous reasons, the depletion region established about the source, at the interface, and at the drain is widest near the drain.

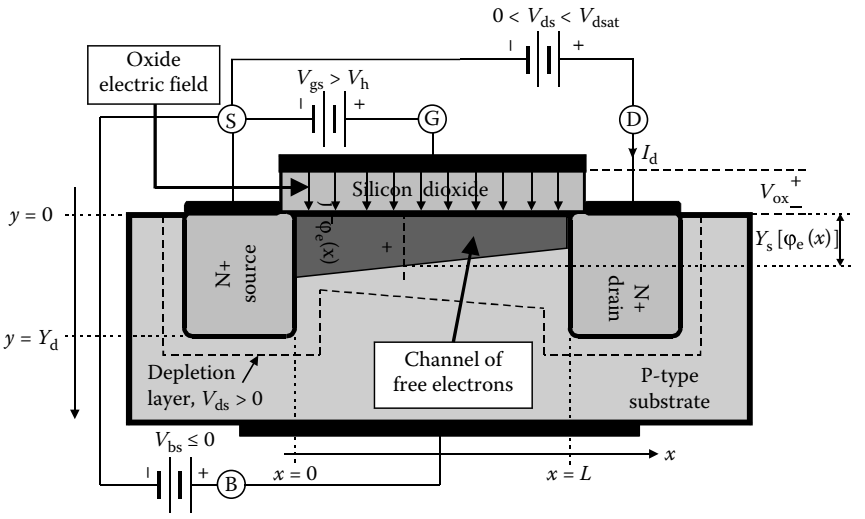


FIGURE 1.26 Cross section of the N-channel MOSFET operated in its ohmic regime. Note that all applied voltages are referred to the source terminal. The diagram is not drawn to scale.

An additionally important point is that the channel potential, symbolized in Figure 1.26 as $\varphi_c(x)$, is measured with respect to the source site. This notation is not to be confused with the previously invoked variable, $\varphi_o(y)$, which measures the potential at the interfacial surface with respect to the neutral region of the bulk substrate. The change in symbolism is reasonable and is encouraged by two issues addressed in the Section 1.2.3.3. The first of these issues is that the bulk–source biasing voltage, V_{bs}, has been absorbed into the threshold voltage metric stipulated by Equation 1.96. Second, this threshold voltage has been defined in terms of the gate voltage, V_{gs}, measured with respect to the source, as opposed to the gate voltage, V_{gb}, referenced to the bulk terminal. Because V_{ds} is nonzero, the channel potential, $\varphi_c(x)$, is not a constant but instead, it varies continuously from $\varphi_c(0) = 0$ at the source site where $x = 0$ to $\varphi_c(L) = V_{ds}$ at the drain site where $x = L$.

An applied drain–source voltage, V_{ds}, launches a lateral electric field, say $E_x(\varphi_c(x))$, that is directed from the drain site-to-the source site and is functionally dependent on the channel potential, $\varphi_c(x)$. This electric field is given by the familiar relationship,

$$E_x(\varphi_c(x)) = -\frac{d\varphi_c(x)}{dx}. \tag{1.100}$$

If μ_n denotes the mobility of electrons, whose concentration within the surface inversion layer postulated in Figures 1.26 and 1.25 is $N_s(\varphi_c(x))$, the static drain current, I_d, promoted by this lateral field is

$$\begin{aligned} I_d &= -q\mu_n W[Y_s(\varphi_c(x))][N_s(\varphi_c(x))]E_x(\varphi_c(x)) \\ &= q\mu_n W[Y_s(\varphi_c(x))][N_s(\varphi_c(x))]\frac{d\varphi_c(x)}{dx}, \end{aligned} \tag{1.101}$$

where $Y_s(\varphi_c(x))$ is the inversion layer thickness introduced in Figure 1.25 and depicted as dependent on the channel potential, $\varphi_c(x)$, in Figure 1.26. It is worthwhile noting that the product, $\mu_n E_x(\varphi_c(x))$, is the velocity of electrons propagated through the inversion layer. This velocity is zero, thereby implying zero drain current, if the gradient, $d\varphi_c(x)/dx$, of channel potential is null. In turn, the channel potential gradient is zero if the applied drain–source voltage, V_{ds}, is zero.

If it is assumed that increases in the channel potential over and above the threshold level incur no change in the geometry of the interfacial depletion region and instead, only cause electrons to be attracted to the surface, Gauss's law predicts

$$q[Y_s(\varphi_c(x))][N_s(\varphi_c(x))] = \varepsilon_{ox}E_{ox}(\varphi_c(x)), \tag{1.102}$$

with $E_{ox}(\varphi_c(x))$ symbolizing the oxide electric field, which is given by

$$E_{ox}(\varphi_c(x)) = \frac{V_{ox}}{T_{ox}} = \frac{V_{gs} - V_h - \varphi_c(x)}{T_{ox}}. \tag{1.103}$$

Recalling Equations 1.63, 1.103, and 1.102 combine with Equation 1.101 to deliver

$$I_d dx = \mu_n C_{ox} W \left[V_{gs} - V_h - \varphi_c(x)\right] d\varphi_c(x). \tag{1.104}$$

An integration of the left-hand side of this result from $x = 0$ to $x = L$ is tantamount to integrating the right-hand side of said result from $\varphi_c(0) = 0$ to $\varphi_c(L) = V_{ds}$. Assuming constant electron mobility through the channel, the requisite integration is straightforward and leads to the desired volt–ampere relationship,

$$I_d = K_n\left(\frac{W}{L}\right)V_{ds}\left(V_{gs} - V_h - \frac{V_{ds}}{2}\right), \tag{1.105}$$

where

$$K_n \triangleq \mu_n C_{ox} \tag{1.106}$$

is the so-called *transconductance coefficient* of the MOSFET. Although K_n is termed a transconductance coefficient, it is not actually a transconductance in that its physical dimension is that of siemens/volt, or mhos/volt.

Several interesting and enlightening features are advanced by Equation 1.105. The first of these is that zero drain current prevails if $V_{ds} = 0$, which is reassuring in that a current flow for null drain–source voltage violates engineering reason, if not the minor issue of conservation of energy. A second, and more significant, point is that the drain current is directly proportional to the gate aspect ratio, W/L. Thus, for fixed gate–source and drain–source voltages, the drain current can be increased or decreased in proportion to this geometric ratio. This controllability over the drain current renders the gate aspect ratio a designable circuit parameter, subject to the proviso that the circuit designer not attempt to make the gate width, W, smaller than the minimum channel length, L, that the process foundry is capable of producing. Thus, if the foundry boasts a 130 nm channel length process, the smallest practical value of W is, in fact, also 130 nm.

A third important feature of Equation 1.105 is the existence of a value of V_{ds} for which drain current I_d is maximized. By setting to zero the partial derivative of I_d in Equation 1.105 with respect to V_{ds}, this extremum is determined to lie at $V_{ds} = (V_{gs} - V_h) = V_{dsat}$, for which the corresponding maximum current, say I_{dsat}, is

$$I_{dsat} = \frac{K_n}{2}\left(\frac{W}{L}\right)\left(V_{gs} - V_h\right)^2 = \frac{K_n}{2}\left(\frac{W}{L}\right)V_{dsat}^2. \tag{1.107}$$

Recalling Equation 1.99, $V_{ds} = V_{dsat}$ corresponds to a gate–drain voltage, V_{gd}, of $V_{gd} = V_h$, which implies that the surface potential at the drain end of the channel barely sustains the onset of strong inversion. In effect, the depth of the electron channel is reduced to zero at the drain site for $V_{ds} = V_{dsat}$,

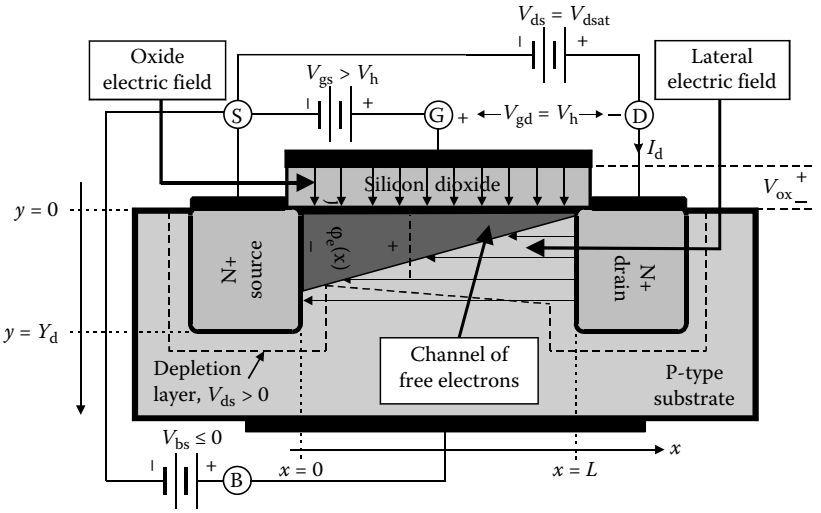


FIGURE 1.27 Cross section of the N-channel MOSFET operated in channel pinch off. The diagram is not drawn to scale.

which justifies the common vernacular of a channel that is *pinched off* at the drain. The situation at hand is diagrammed in Figure 1.27.

At first blush, it may appear incongruous that a pinched off channel, which might be viewed as a means to cut off the supply of electrons to the drain site, can sustain a drain current, yet alone the maximum drain current postulated by Equation 1.107. The current is indeed sustained because of two prevailing phenomena. First, the electric field, $-\mathrm{d}\varphi_c(x)/\mathrm{d}x$, within the inversion layer encourages the transit of electrons toward the tapered edge of the channel at the drain site. Second, electrons reaching the channel edge are influenced immediately by the lateral electric field established by the applied positive drain-to-source voltage. This field, which is abstracted in Figure 1.27 by the indicated horizontal vectors directed from the drain region-to-the source region, sweeps those electrons at the inversion layer boundary into the drain region. The resultant current arising from the transport of electrons across the depletion zone between the tapered channel edge and the drain is, like the current within the inversion layer, proportional to the mobility of electrons. In the depletion zone, this carrier mobility is minority carrier mobility, which is inversely proportional to the background impurity concentration of the bulk substrate. A fundamental reason for maintaining relatively low impurity concentration in the bulk is the assurance of relatively high minority carrier (electron) mobility therein so that carriers are swept across the depletion zone at high velocity, thereby facilitating fast transistor switching and broadband circuit responses.

The fourth interesting feature surrounding Equation 1.105 lends credence to the term, "ohmic," as a descriptive for the operating regime at hand. In particular, Equation 1.105 can be expressed in the form,

$$I_d = K_n\left(\frac{W}{L}\right)V_{ds}\left(V_{gs} - V_h - \frac{V_{ds}}{2}\right) = \frac{V_{ds}}{R_{ds}(V_{gs})}, \tag{1.108}$$

with

$$R_{ds}(V_{gs}) = \frac{1}{K_n(W/L)\left(V_{gs} - V_h - \frac{V_{ds}}{2}\right)}. \tag{1.109}$$

In other words, and as is proffered in Figure 1.28, a MOSFET operated in its ohmic regime, where $V_{gs} \ge V_h$ and $V_{ds} \le V_{dsat}$, behaves as a drain-to-source resistance, $R_{ds}(V_{gs})$, whose resistance value is

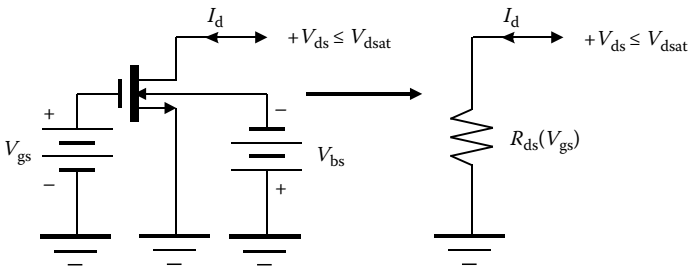


FIGURE 1.28 Static circuit model of an N-channel MOSFET operated in its ohmic regime. The transistor can be operated in such a way that its drain–source terminals emulate a voltage-controlled, nominally linear resistance.

controlled by the applied gate–source voltage, V_{gs}. Moreover, the synthesized resistance is approximately independent of the voltage, V_{ds}, developed across its terminals, and therefore emulates a linear resistance, if $V_{ds} \ll 2(V_{gs} - V_h) \equiv 2V_{dsat}$. In effect, the ohmic regime MOSFET is an electronic approximation of a linear potentiometer whose resistance setting is inversely proportional to the applied gate–source voltage.

1.2.3.3 Saturation Regime

In saturation, which is the volt–ampere domain in which MOSFETs embedded in high-performance analog circuits function, $V_{gs} \ge V_h$ and $V_{ds} \ge (V_{gs} - V_h)$. To first order, the drain current in saturation is taken to be the drain saturation current given by Equation 1.107, which is independent of drain–source voltage, V_{ds}, that is,

$$I_d = \frac{K_n}{2}\left(\frac{W}{L}\right)\left(V_{gs} - V_h\right)^2 \quad \text{for } V_{gs} > V_h, \quad \text{and} \quad V_{ds} \ge \left(V_{gs} - V_h\right). \tag{1.110}$$

The logic underlying this approximation is that the drain current in saturation is determined by the surface electron concentration established for the drain–source voltage, $V_{ds} = (V_{gs} - V_h) = V_{dsat}$, which barely allows for an electron channel spanning the entire source-to-drain spacing. Any increase in the drain-to-source voltage above its saturated value, V_{dsat}, simply adds impetus to the attractive force exerted on inversion layer electrons by the lateral electric field promoted by the drain–source voltage.

The problem with the foregoing logic is that the drain current given by Equation 1.107 is premised on Equation 1.105, which in turn invokes the presumption of an electron inversion layer length that is identical to the channel spacing length, L, separating the source region from the drain region. If $V_{ds} = V_{dsat}$ incurs pinch off at the drain site, and hence an inversion layer length equal to the channel length, L, as illustrated in Figures 1.27 and 1.29a, $V_{ds} > V_{dsat}$ necessarily incurs pinch off within the source–drain spacing, as is suggested in Figure 1.29b. Because of the indicated reduction in the effective channel length from L to $(L - \Delta L)$, the integrated form of Equation 1.104 is now

$$\int_0^{L-\Delta L} I_d dx = \int_0^{V_{dsat}} \mu_n C_{ox} W\left[V_{gs} - V_h - \varphi_c(x)\right] d\varphi_c(x). \tag{1.111}$$

The result of this integration exercise is easily demonstrated to be

$$I_d = \frac{K_n}{2}\left(\frac{W}{L - \Delta L}\right)\left(V_{gs} - V_h\right)^2 = I_{dsat}\left(\frac{L}{L - \Delta L}\right), \tag{1.112}$$

where the current, I_{dsat}, is given by Equation 1.107 and represents the drain current at the transition boundary between ohmic and saturation operational regimes. For most practical applications of

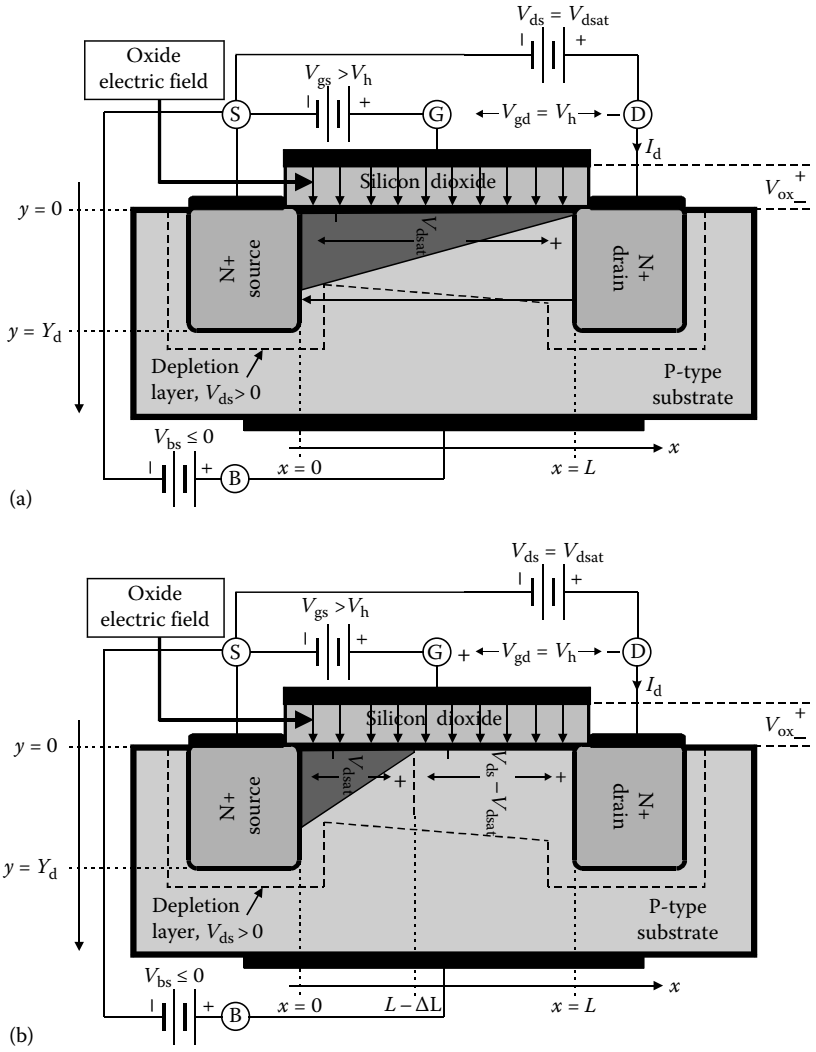


FIGURE 1.29 (a) Cross section of N-channel MOSFET operated in strong inversion and with $V_{ds} = V_{dsat}$. (b) Cross section of the MOSFET in (a) operated with $V_{ds} > V_{dsat}$. The diagrams are not drawn to scale.

MOSFETs [8], the effective reduction, $(L - \Delta L)$, in channel length relates to the drawn channel length, L, as

$$\frac{L}{L - \Delta L} \approx 1 + \frac{V_{ds} - V_{dsat}}{V_\lambda}, \tag{1.113}$$

where V_λ, termed the *channel length modulation voltage*,* is given by the semiempirical expression,

$$V_\lambda = \left(\frac{L}{D_b}\right)\left(\frac{V_j}{V_F}\right)^2 \sqrt{32 V_T \left(V_{ds} - V_{dsat} + V_j\right)}. \tag{1.114}$$

* Most HSPICE and other SPICE simulators use a *channel length parameter* to compute the degree to which the drain–source voltage affects the drain saturation current. This channel length parameter, λ, is $\lambda = 1/V_\lambda$.

In Equation 1.114, V_T is the familiar thermal voltage, D_b is the electron Debye length delineated in Equation 1.56, and V_j is the *built-in potential* of the bulk–drain PN junction. Specifically,

$$V_j = V_T \ln\left(\frac{N_A N_D}{N_i^2}\right), \tag{1.115}$$

with N_A, N_D, and N_i respectively denoting the average impurity concentration in the bulk substrate, the average impurity concentration of the drain diffusion (or implant), and the intrinsic carrier concentration of silicon. Equation 1.114 delivers acceptable analytical accuracy for channel lengths, L, that are no smaller than 0.09 μ and drain–source voltages, V_{ds}, that lie within breakdown ratings of the considered transistor.

The drain current in the saturation regime is now expressible as

$$I_d \approx \frac{K_n}{2}\left(\frac{W}{L}\right)(V_{gs} - V_h)^2\left(1 + \frac{V_{ds} - V_{dsat}}{V_\lambda}\right) = I_{dsat}\left(1 + \frac{V_{ds} - V_{dsat}}{V_\lambda}\right), \tag{1.116}$$

where it is understood that the gate–source and drain–source voltages, V_{gs} and V_{ds}, respectively, are constrained to satisfy the saturation requirements, $V_{gs} > V_h$ and $V_{ds} \geq (V_{gs} - V_h) = V_{dsat}$. Clearly, the saturation regime drain current is no longer independent of the drain–source voltage. The current is seen to rise with V_{ds} with a slope of I_{dsat}/V_λ. Note, however, that this slope is not constant owing to its square root dependence on V_{ds}. For large V_λ, which is manifested by long transistor channel length, L, this rate of current rise with V_{ds} is modest and indeed, the slope of the current–voltage characteristic curve approaches zero in the limit as V_λ approaches infinity. These observations and Equation 1.116 itself suggest that the drain–source port of a MOSFET does not behave as a constant current source whose value, I_{dsat}, is controlled exclusively by gate–source voltage V_{gs}. Instead, the drain–source port is a practical controlled current source comprised of a constant current generator, albeit controlled by gate–source voltage V_{gs}, in shunt with a resistive branch. To wit, Equation 1.116 can be written as

$$I_d \approx I_{dsat} + \frac{V_{ds} - V_{dsat}}{V_\lambda / I_{dsat}}, \tag{1.117}$$

which suggests the static circuit model provided in Figure 1.30. The subject model is more useful conceptually than computationally since a change made to V_{gs} for the purpose of adjusting the nominal drain current, I_{dsat}, influences the resistance value, V_λ/I_{dsat}, and the voltage offset, V_{dsat}, introduced in the drain–source port.

A complication of the channel length embellishment to the saturation drain current expression is that Equation 1.116 is discontinuous with the ohmic domain drain current expression in Equation 1.105 at the transition boundary between respective operating domains. Simple software fixes in commonly

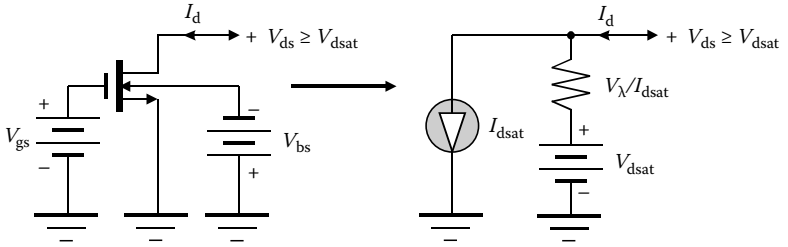


FIGURE 1.30 A large-signal circuit model for an N-channel MOSFET biased to operate in its saturation domain.

available circuit simulators rectify this incongruity. From an analytical perspective, the problem can be tacitly ignored, if V_λ in Equation 1.114 abides by the previously disclosed channel length and voltage restrictions.

For the convenience of the reader, the relevant expressions for the volt–ampere characteristic curves of an N-channel MOSFET are synopsized herewith. In particular,

$$I_d \approx \begin{cases} 0, & V_{gs} < V_h \\ K_n\left(\frac{W}{L}\right) V_{ds}\left(V_{gs} - V_h - \frac{V_{ds}}{2}\right), & V_{gs} \geq V_h;\ V_{ds} < V_{dsat} \\ \frac{K_n}{2}\left(\frac{W}{L}\right)\left(V_{gs} - V_h\right)^2\left(1 + \frac{V_{ds} - V_{dsat}}{V_\lambda}\right), & V_{gs} \geq V_h;\ V_{ds} \geq V_{dsat} \end{cases}, \tag{1.118}$$

where V_{dsat} is the voltage difference, $(V_{gs} - V_h)$. It is to be understood that the positive reference direction of the drain current in NMOS is a current flowing into the drain, while the positive reference voltage polarities reflect those highlighted in Figure 1.17. Moreover, V_h is recalled as a threshold level dependent on the bulk–source voltage, V_{bs}, in accordance with Equation 1.96. A representative plot of the static volt–ampere characteristics of an NMOS transistor appear in Figure 1.31.

In the interests of clarity and completeness, the PMOS counterpart to Equation 1.118 is

$$I_d \approx \begin{cases} 0, & V_{sg} < V_h \\ K_p\left(\frac{W}{L}\right) V_{sd}\left(V_{sg} - V_h - \frac{V_{sd}}{2}\right), & V_{sg} \geq V_h;\ V_{sd} < V_{dsat} \\ \frac{K_p}{2}\left(\frac{W}{L}\right)\left(V_{sg} - V_h\right)^2\left(1 + \frac{V_{sg} - V_{dsat}}{V_\lambda}\right), & V_{sg} \geq V_h;\ V_{sd} \geq V_{dsat} \end{cases}, \tag{1.119}$$

where, in terms of the source–gate voltage, V_{sg}, V_{dsat} is now given by, $(V_{sg} - V_h)$, the threshold voltage, which is dependent on source–bulk voltage V_{sb} in Equation 1.96, remains a positive number, and

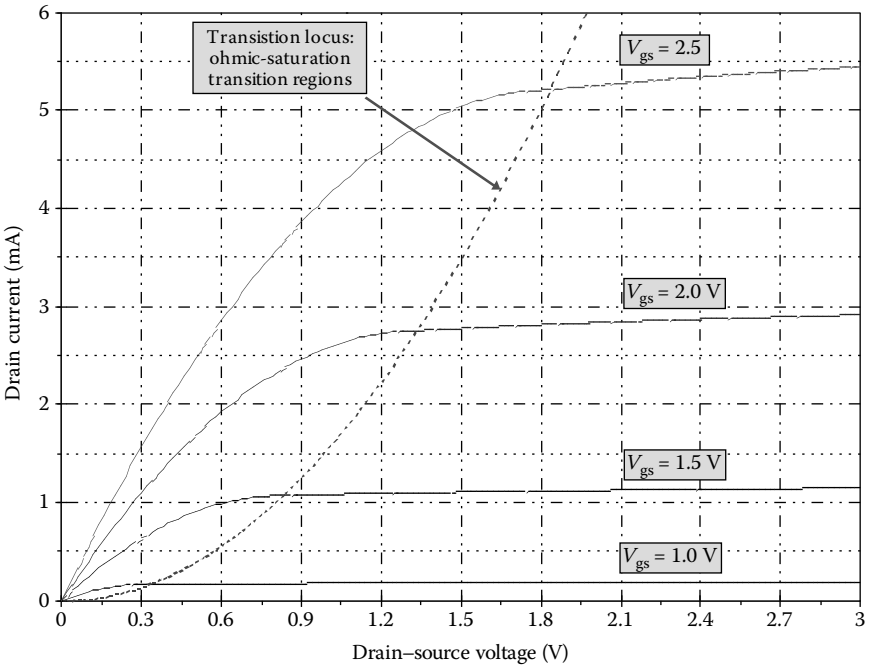


FIGURE 1.31 Common-source volt–ampere characteristic curves for an NMOS transistor.

transconductance parameter K_p is now the product of oxide capacitance density and hole mobility. The positive reference direction of the drain current in PMOS is a current flowing out of the drain, while the positive reference voltage polarities pertain to those delineated in Figure 1.18.

1.2.3.4 Refinements to the Static Model

The static volt–ampere characteristic in Equations 1.118 and 1.119 exhibit observable errors when computed currents are compared to experimental measurements executed on deep submicron MOSFET technology transistors. The principle source of these errors is two types of mobility degradation to which carriers in the source-to-drain channel are subjected. The first form of mobility degradation derives from the large lateral electric fields evidenced when even relatively small drain-to-source voltages are applied across channels whose lengths are smaller than approximately 0.25 μ. The second form of mobility impairment is caused by the strong vertical electric fields established by gate–source voltages applied across thin oxide layers.

1.2.3.4.1 Lateral Electric Fields

The NMOS and PMOS volt–ampere characteristic equations in Equations 1.118 and 1.119 are predicated on the presumption that the drift velocity, say v_c, of carriers propagated through the inverted channel at the oxide–semiconductor interface is proportional to the lateral electric field, $E_x[\varphi_c(x)]$. This field is, of course, established in the channel by applied drain–source voltage, V_{ds}, (in the case of NMOS) or applied source–drain voltage V_{sd} (in the case of PMOS). In particular,

$$v_c = \mu_o |E_x|, \tag{1.120}$$

where μ_o represents either the low field value of the electron mobility, μ_n, in N-channel devices or the low field value of the hole mobility, μ_p, in PMOS. The simpler notation, E_x is adopted in Equation 1.120 to represent the potential-dependent field function, $E_x[\varphi_c(x)]$. The need for the absolute value operation on the right-hand side of Equation 1.120 materializes from the fact that the carrier velocity, which is always a positive metric, is directed against the direction of the channel field in NMOS. In the case of NMOS transistors, carriers drift in the direction of the source-to-the drain, whereas the field is directed from drain-to-source and is therefore negative. For PMOS, no algebraic sign problems are manifested, since carriers drift in the same direction as the lateral field, whence E_x is positive.

The simplicity of Equation 1.120 belies the fact that the carrier drift velocity does not continually increase in proportion to the electric field. In fact, the carrier velocity saturates at a value, say v_{max}, which is of the order of 0.15 μm/ps in silicon, when electric fields are excessive. In recognition of this physical phenomenon, Equation 1.120 is supplanted by the empirical relationship,

$$v_c = \frac{\mu_o |E_x|}{1 + |E_x|/E_c}, \tag{1.121}$$

where

$$E_c = \frac{v_{max}}{\mu_o} \tag{1.122}$$

is termed the *critical electric field*. Typically, E_c is in the range of 3–5 V/μm. A comparison of Equation 1.121 with Equation 1.120 suggests an effective mobility, μ_e, of

$$\mu_e = \frac{\mu_o}{1 + |E_x|/E_c}. \tag{1.123}$$

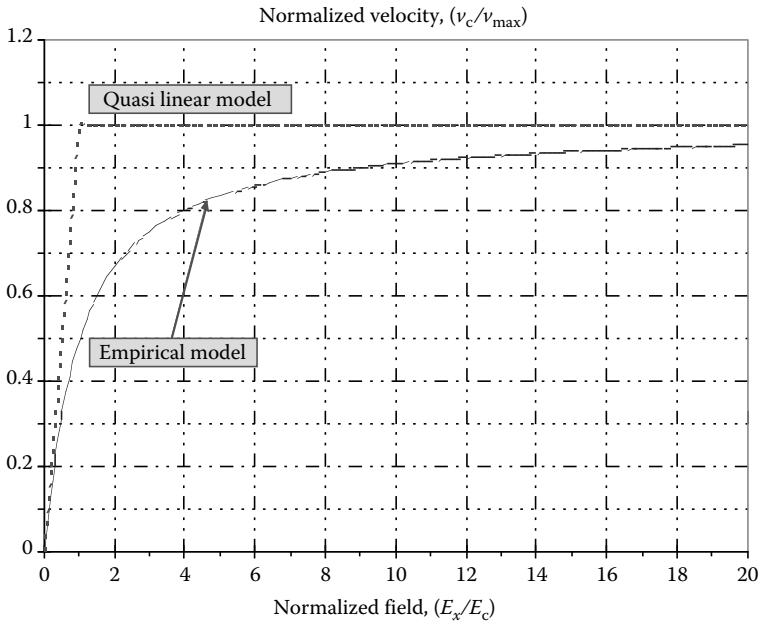


FIGURE 1.32 The dependence of carrier velocity on electric field in a semiconductor. The dashed curve represents the elementary low field approximation to the velocity-field relationship.

The mobility degradation inferred by the last disclosure bodes potentially decreased frequency response attributes of considered transistors, since the less mobile free electrons are in the inverted channel, the longer is the average time required for their transport from the source-to-the drain. Figure 1.32 sketches the velocity-field relationship implied by Equation 1.121. Note in this plot that the linear, or low field, approximation to the velocity characteristic is reasonably accurate up to about only 30% of the saturated limited velocity.

The effect on the ohmic regime drain current of the mobility degradation incurred by strong lateral electric fields can be studied by returning to Equation 1.104 and replacing the electron mobility, μ_n, therein by an adjusted mobility, μ_{ne}, such that

$$\mu_{ne} = \frac{\mu_n}{1 - E_x/E_c} = \frac{\mu_n}{1 + \frac{1}{E_c}\frac{d\varphi_c(x)}{dx}}, \tag{1.124}$$

where Equation 1.100 is applied and μ_n is understood to be the low field value of electron mobility in the inverted source-to-drain channel. Equation 1.104 becomes

$$I_d = \frac{\mu_n C_{ox} W\left[V_{gs} - V_h - \varphi_c(x)\right]}{1 + \frac{1}{E_c}\frac{d\varphi_c(x)}{dx}}\frac{d\varphi_c(x)}{dx}, \tag{1.125}$$

which leads to

$$I_d\left[\int_0^L dx + \frac{1}{E_c}\int_0^{V_{ds}} d\varphi_c(x)\right] = \mu_n C_{ox} W \int_0^{V_{ds}} \left[V_{gs} - V_h - \varphi_c(x)\right] d\varphi_c(x). \tag{1.126}$$

The requisite integrations produce

$$I_{\mathrm{d}} = K_n\left(\frac{W}{L}\right)\left[\frac{V_{\mathrm{ds}}\left(V_{\mathrm{gs}} - V_{\mathrm{h}} - \frac{V_{\mathrm{ds}}}{2}\right)}{1 + \frac{V_{\mathrm{ds}}}{V_{\mathrm{le}}}}\right], \tag{1.127}$$

where

$$V_{\mathrm{le}} = E_{\mathrm{c}}L = \left(\frac{v_{\max}}{\mu_n}\right)L \tag{1.128}$$

might be termed the *lateral electric field modulation voltage*. Observe that Equation 1.127 differs from the ohmic region volt–ampere relationship in Equation 1.118 by only the dimensionless factor in the denominator on the right-hand side of Equation 1.127. Appealing to Equation 1.128, this factor is seen to approach one when the channel length, L, is long. Of course, the subject factor also tends toward unity if the drain–source voltage, V_{ds}, is small. The latter point reflects engineering expectations in that small V_{ds} incurs lateral electric fields that are small enough to minimize field-induced mobility degradation.

A complication spawned by Equation 1.127 is that it no longer delivers the simple relationship for the drain saturation voltage witnessed in Equation 1.99. By definition, the drain saturation voltage, V_{dsat}, is the value of the drain–source voltage, V_{ds}, for which the slope of the ohmic regime I_{d} versus V_{ds} characteristic is zero. An application of this definition to Equation 1.127 leads to the revised drain saturation voltage,

$$V_{\mathrm{dsat}} = M_{\mathrm{sat}}\left(V_{\mathrm{gs}} - V_{\mathrm{h}}\right), \tag{1.129}$$

where, with

$$\alpha \triangleq \frac{V_{\mathrm{gs}} - V_{\mathrm{h}}}{V_{\mathrm{le}}}, \tag{1.130}$$

$$M_{\mathrm{sat}} = \frac{\sqrt{1 + 2\alpha} - 1}{\alpha}. \tag{1.131}$$

It can be demonstrated that $M_{\mathrm{sat}} \leq 1$ for $\alpha \geq 0$ and thus, an impact of carrier mobility degradation incurred by strong lateral fields in the inverted channel is a decrease in the low field value of the drain saturation voltage. While mobility degradation is generally an undesirable phenomenon, the drain saturation voltage decrease is actually good news in low-voltage applications that require MOSFETs to function in their saturated regimes.

The drain saturation current corresponding to the revised estimate of the drain saturation voltage can be determined by substituting Equation 1.129 into 1.127. This activity produces the aesthetically pleasing result,

$$I_{\mathrm{dsat}} = \frac{K_n}{2}\left(\frac{W}{L}\right)V_{\mathrm{dsat}}^2 = \frac{K_n}{2}\left(\frac{W}{L}\right)M_{\mathrm{sat}}^2\left(V_{\mathrm{gs}} - V_{\mathrm{h}}\right)^2. \tag{1.132}$$

In the limit of large channel lengths, V_{le} in Equation 1.128 is large, thereby rendering parameter α in Equation 1.130 small. But for very small α, M_{sat} in Equation 1.131 approaches unity. It is therefore reassuring that in the limit of large channel lengths, which are incapable of supporting large electric fields in the inverted channel, I_{dsat} in Equation 1.132 collapses to Equation 1.107, a relationship that implicitly

reflects tacit neglect of field-induced carrier mobility degradation. In contrast, very small channel lengths give rise to small V_{le} and large α, whence M_{sat} in Equation 1.131 reduces to

$$M_{sat}\big|_{\text{small L}} = \frac{\sqrt{1+2\alpha}-1}{\alpha}\bigg|_{\text{large }\alpha} \approx \sqrt{\frac{2}{\alpha}}. \tag{1.133}$$

Upon combining the last result with Equation 1.132, the short channel value of I_{dsat} is found to be

$$I_{dsat}\big|_{\text{small L}} \approx WC_{ox}v_{max}\left(V_{gs}-V_h\right), \tag{1.134}$$

where Equations 1.128 and 1.106 are exploited. Observe that the resultant drain saturation current is independent of the channel length, L. This independence stems from the fact that in the limit of very small channel lengths, carriers (electrons in the present case of an NMOS transistor) are transported through the inverted channel at their saturated limited, or maximum, velocity. This maximum velocity of carrier propagation renders L inconsequential with respect to the average time of carrier transport from the source region-to-the drain region. But perhaps the most interesting aspect of Equation 1.134 is that the short channel drain saturation current is a linear function of the gate–source voltage, V_{gs}. The linearity posed by Equation 1.134 is an obvious advantage for most analog signal processing applications, but achieving the velocity saturation implicit to this observed linearity may present voltage biasing challenges.

Of course, Equation 1.129 through 1.133 apply to the saturation regime of device operation in that in saturation, the drain current is merely the transistor current, I_{dsat}, evidenced at the boundary of ohmic and saturation regimes, corrected by channel length modulation effects. To wit, short channel phenomena imply that for $V_{gs} \geq V_h$ and $V_{ds} \geq V_{dsat}$,

$$I_d = \frac{K_n}{2}\left(\frac{W}{L}\right)M_{sat}^2\left(V_{gs}-V_h\right)^2\left(1+\frac{V_{ds}-V_{dsat}}{V_\lambda}\right), \tag{1.135}$$

where it is essential to remember that the drain saturation voltage, V_{dsat}, is now given by Equation 1.129. It is clear that M_{sat} in Equation 1.129 is properly viewed as a drain saturation voltage correction factor in a short channel (indeed, deep submicron) environment. Because of Equation 1.135, the square of M_{sat} can be accorded the stature of a current correction factor pertinent to short channel drain currents in the saturation regime. The dependence on parameter α of these correction factors is displayed in the plots submitted in Figure 1.33. The indicated correction factors are significant. For example, consider $\alpha = 2$, which might typically represent a gate–source voltage, V_{gs}, that is about a volt over the threshold potential. The curves in the figure at hand suggest an approximate 38% reduction in the drain saturation voltage predicted by the simple long channel model, which corresponds to $\alpha = 0$, as well as about a 62% attenuation of the corresponding drain saturation current.

Although Equations 1.135 and 1.129 are analytically elegant, their utility in a design-oriented environment is questionable in light of the dependence of factor M_{sat} on parameter α set forth by Equation 1.131. In light of this dilemma, an approximate curve fit of both M_{sat} and its square is judicious from an engineering design perspective. A numerical study of Equation 1.131 reveals that the empirical approximation,

$$M_{sat} = \frac{\sqrt{1+2\alpha}-1}{\alpha} \approx 1-\frac{\sqrt{\alpha}}{4}, \tag{1.136}$$

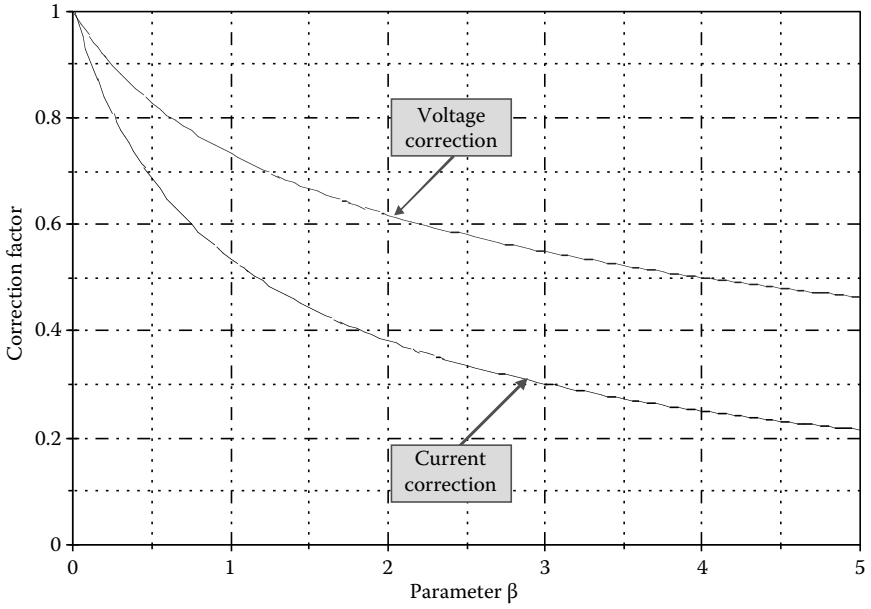


FIGURE 1.33 Voltage and current correction factors precipitated by large lateral electric fields in short channel MOSFETs. The parameter, α, is the effective gate–source voltage, $(V_{gs} - V_h)$, normalized to the lateral electric field modulation voltage, V_{le}.

results in an error of at most 4.8% for $0 \leq \alpha \leq 5$. On the other hand, a similar numerical exercise produces

$$M_{sat}^2 = \left(\frac{\sqrt{1+2\alpha}-1}{\alpha}\right)^2 \approx \frac{1}{1+0.78\alpha} \tag{1.137}$$

to a computational error of at most 5.1% for $0 \leq \alpha \leq 5$. For most design-oriented purposes, Equation 1.129 can therefore be supplanted by

$$\begin{aligned} V_{dsat} = M_{sat}(V_{gs} - V_h) &\approx \left(1 - \frac{\sqrt{\alpha}}{4}\right)(V_{gs} - V_h) \\ &= \left(1 - \frac{1}{4}\sqrt{\frac{V_{gs} - V_h}{V_{le}}}\right)(V_{gs} - V_h), \end{aligned} \tag{1.138}$$

while Equation 1.135 becomes for circuit design applications of MOSFETs operated in saturated regimes,

$$\begin{aligned} I_d &= \frac{K_n}{2}\left(\frac{W}{L}\right)M_{sat}^2(V_{gs} - V_h)^2\left(1 + \frac{V_{ds} - V_{dsat}}{V_\lambda}\right) \\ &\approx \frac{K_n}{2}\left(\frac{W}{L}\right)(V_{gs} - V_h)^2\left(\frac{1 + \frac{V_{ds} - V_{dsat}}{V_\lambda}}{1 + 0.78\alpha}\right) \\ &= \frac{K_n}{2}\left(\frac{W}{L}\right)(V_{gs} - V_h)^2\left(\frac{1 + \frac{V_{ds} - V_{dsat}}{V_\lambda}}{1 + 0.78\left(\frac{V_{gs} - V_h}{V_{le}}\right)}\right). \end{aligned} \tag{1.139}$$

The academic purist who may understandably balk at the foregoing numerical empiricisms is respectfully reminded that the mobility expression in Equation 1.123 and the ''long channel'' velocity relationship of Equation 1.120 are hardly grounded in sound physical phenomenology. Moreover, it is interesting to note that of the more than 275 parameters indigenous to the commonly exploited Level 49 HSPICE model of a MOSFET, most are curve fit disclosures that bear no clarion relationship to the physical charge storage and charge transport mechanisms that underpin the volt–ampere characteristics of a MOSFET.

1.2.3.4.2 Vertical Electric Fields

Apart from the carrier mobility degradation incurred by strong lateral fields in the inverted channel of a MOSFET, mobility is impacted by the vertical electric field resulting from the applied effective interface potential, $(V_{gs} - V_h)$, in the case of NMOS or $(V_{sg} - V_h)$ for PMOS devices. In NMOS, increases in V_{gs} strengthens this vertical field so that free electrons transported from the source-to-the drain are encouraged to drift ever closer to the oxide–semiconductor interface. Unfortunately, the interface is far from a perfectly smooth boundary, if for no other reason than routine device processing invariably produces ionic contamination therein. The imperfect boundary causes potentially significant carrier scattering, which in turn results in diminished carrier mobility.

To first order, the mobility attenuation resulting from increased gate overdrive can be addressed analytically by replacing the low field mobility, μ_n (for NMOS), to which K_n in Equation 1.139 is directly proportional, by an effective carrier mobility, μ_{eff}, such that

$$\mu_{eff} = \frac{\mu_n}{1 + \frac{V_{gs} - V_h}{V_{ve}}}. \tag{1.140}$$

In this expression, V_{ve} is the *vertical electric field modulation voltage*, which is nominally directly proportional to the thickness, T_{ox}, of the oxide layer. Of course, an expression analogous to Equation 1.140 prevails for hole mobility in the inverted channel of PMOS transistors. To a very rough approximation,

$$V_{ve} = T_{ox}/15, \tag{1.141}$$

where T_{ox} in units of angstroms returns V_{ve} in units of volts. Because of Equations 1.140 and 1.139 for the saturation domain current becomes

$$\begin{aligned} I_d &= \frac{K_n}{2}\left(\frac{W}{L}\right) M_{sat}^2 \frac{\left(V_{gs} - V_h\right)^2}{\left(1 + \frac{V_{gs} - V_h}{V_{ve}}\right)} \left(1 + \frac{V_{ds} - V_{dsat}}{V_\lambda}\right) \\ &\approx \frac{K_n}{2}\left(\frac{W}{L}\right) \frac{\left(V_{gs} - V_h\right)^2}{\left(1 + \frac{V_{gs} - V_h}{V_{ve}}\right)} \left(\frac{1 + \frac{V_{ds} - V_{dsat}}{V_\lambda}}{1 + 0.78\left(\frac{V_{gs} - V_h}{V_{le}}\right)}\right). \end{aligned} \tag{1.142}$$

An analogous modification, which amounts to an effective reduction of the transconductance parameter, K_n, can be made to the ohmic domain current.

Obviously, Equation 1.142 is inordinately more cumbersome than is the simple, square law, volt–ampere characteristic advanced by Equation 1.110 for device operation in the saturation domain. As a result, the design-oriented determination of a suitable gate–source voltage for a desired drain current and corresponding drain–source voltage can be a daunting challenge. But in addition to the computational problems precipitated merely by algebraic complexity, engineering difficulties are additionally encountered with respect to the accurate numerical delineation of the model metrics, K_n, V_h, V_{ve}, V_λ, and V_{le}. These

latter difficulties derive from the unfortunate fact that the physical device and charge transport properties (saturation velocity, carrier mobility, regional concentrations, etc.) on which these and other model parameters depend are invariably unavailable to the circuit designer. At best, the circuit designer can reasonably expect to have presumably reliable, detailed device model parameters suitable for computer-aided simulation of transistor performance. For example, process foundries routinely supply their customers with device models in the form of Level 49 HSPICE or other computer-based files. Unfortunately, many, if not most, of the hundreds of numerical entries indigenous to these files are themselves nonphysical entities that defy satisfying mathematical relationships to the physical model metrics discussed in earlier paragraphs. These and related other design-oriented problems can prove exasperating. The aforementioned issues are best mitigated by coalescing manual design strategies and calculations with suitable computer-based simulations of device properties and volt–ampere characteristics.

1.2.3.5 Temperature Effects

The operating temperature of the inverted interfacial channel affects the drain current of a transistor in three ways. First, because thermal energy imparted to free carriers increases their scattering, the carrier mobility decreases in response to increased operating temperatures. To first order, the electron mobility, $\mu_n(T)$, at absolute temperature T relates to the mobility, $\mu_n(T_o)$, at a reference temperature, T_o, in accordance with the three-halves power law,

$$\mu_n(T) = \mu_n(T_o)\left(\frac{T_o}{T}\right)^{3/2}. \tag{1.143}$$

Because parameter K_n in Equation 1.142 is directly proportional to carrier mobility, Equation 1.143 implies that the drain current of a MOSFET is characterized by a negative temperature coefficient, that is, the drain current, I_d, decreases with increasing operating temperature.

A second effect of increased thermal energy is a perturbation of threshold voltage. A computation of this perturbation is best initiated by returning to Equation 1.96 to evaluate the derivative of the threshold voltage, V_h, with respect to the Fermi potential, V_F. Recalling Equations 1.96 and 1.97, and noting that the body effect voltage, V_θ, in Equation 1.90 is independent of temperature,

$$\frac{dV_h}{dV_F} = 2 + \sqrt{\frac{V_\theta}{V_F}} + \frac{V_h - V_{ho}}{2V_F} + \left(\frac{V_\theta}{2V_F}\right)\left(\frac{V_{bs}}{V_h - V_{ho} + 2\sqrt{V_\theta V_F}}\right), \tag{1.144}$$

where V_{ho} is recalled as signifying the zero bias ($V_{bs} = 0$) value of the threshold potential. Note that the last two terms on the right-hand side of this expression vanish when a MOSFET is operated with $V_{bs} = 0$. The sensitivity of the threshold voltage with respect to temperature follows as

$$\frac{dV_h}{dT} = \frac{dV_h}{dV_F} \times \frac{dV_F}{dT}. \tag{1.145}$$

The temperature derivative of the Fermi potential derives from Equation 1.55, with the proviso that due account be made of the temperature dependence of the intrinsic carrier concentration, N_i. To this end, a commonly used empiricism is

$$N_i = N_{io} 2^{(T - T_o)/T_n}, \tag{1.146}$$

where T_n is generally taken to be 10°C and, assuming the reference temperature, T_o, is 27°C, N_{io}, the intrinsic carrier concentration at $T = T_o$, is the previously used number, $(1.45)(10^{10})$ atoms/°C. With

$T_n = 10°C$, Equation 1.146 allows N_i to double for each 10°C rise above the reference temperature. Armed with Equations 1.146 and 1.55 produces

$$\frac{dV_F}{dT} = \frac{V_F}{T} - \frac{V_T}{T_n} \ln 2. \tag{1.147}$$

Equations 1.144 and 1.147 combine to yield the final result,

$$\frac{dV_h}{dT} = \left[2 + \sqrt{\frac{V_\theta}{V_F}} + \frac{V_h - V_{ho}}{2V_F} + \left(\frac{V_\theta}{2V_F}\right)\left(\frac{V_{bs}}{V_h - V_{ho} + 2\sqrt{V_\theta V_F}}\right)\right]\left(\frac{V_F}{T} - \frac{V_T}{T_n} \ln 2\right), \tag{1.148}$$

where parameters V_{ho}, V_h, and V_T, are computed at the reference temperature, T_o. The indicated temperature derivative of the threshold voltage is invariably a positive number in the range of 1.5–2.4 mV/°C. Thus, the threshold voltage increases with increasing operating temperatures, thereby leading to a decrease in the drain current. In other words, the temperature dependence of both the carrier mobility and the threshold voltage conduce a drain current exuding a negative temperature coefficient.

The algebraic form of Equation 1.148 is thoroughly depressing and is hardly a relationship stored in the human memories of circuit designers. Fortunately, for MOSFETs featuring thin gate oxides (under 50 Å) and substrate doping concentrations no smaller than 10^{15} atoms/cm^3, the terms in V_θ, $(V_h - V_{ho})$, and V_T/T_n are generally negligible, especially if the bulk–source bias, V_{bs}, is no more negative than 1.5 V. In this event,

$$\frac{dV_h}{dT} \approx \frac{2V_F}{T}, \tag{1.149}$$

which can be shown to be always larger—generally by no more than 5% or 6%—than the result predicted by Equation 1.148. For a substrate doping concentration of $N_A = (5)(10^{15})$ atoms/cm^3, Equation 1.149 predicts a threshold voltage sensitivity at $T = 27°C = 300.16$ K of 2.2 mV/°C.

1.2.4 Transistor Capacitances

At this juncture, the volt–ampere characteristic equations given by Equations 1.118, 1.119, 1.139, and 1.142 pertain to MOSFETs operated exclusively under static or low-frequency signal conditions. Specifically, the drain currents predicted by these relationships are unrealistically cavalier in that they respond instantaneously to applied gate–source, drain–source, and bulk–source excitations. When high-frequency signals are applied, the current responses are slowed by device capacitances arising from the charge storage that prevails in the inverted channel and within the depletion regions formed about the source and drain diffusions or implants. The engineering implications of this inherent inability of drain currents to respond instantly to signal excitations are MOSFET circuits exuding constrained bandwidths, nonzero input/output (I/O) delays and phase shifts, and nonzero rise and fall times in transient responses. In extreme cases, the interaction of these device capacitances with the energy storage elements of the peripheral circuit can produce excessive response peaking in either the frequency or time domains and even outright instability.

1.2.4.1 Depletion Capacitances

The first of the two principle sources of transistor capacitances is the depletion capacitance indigenous to both of the PN junctions formed respectively between the bulk and drain and between the bulk and source. In turn, each of these two transition region capacitances consists of a planar component and a peripheral, or sidewall, component. The planar component embodies the depletion layer established

between the bulk substrate region and the underside of the source and drain regions. On the other hand, the sidewall capacitance embraces the depletion layers in the areas of the source and drain regions that are proximate to the front surface, the back surface, and the side surface area adjacent to the active channel region. For the bulk–drain depletion capacitance, C_{bd},

$$C_{bd} = \frac{A_d C_j}{\left(1 - \frac{V_{bd}}{V_j}\right)^{M_j}} + \frac{P_d C_{jsw}}{\left(1 - \frac{V_{bd}}{V_j}\right)^{M_{jsw}}}, \tag{1.150}$$

where the forms of each of the terms on the right-hand side are observed to mirror the traditional depletion capacitance associated with a back biased PN junction. In Equation 1.150, C_j is the zero bias (meaning, $V_{bd} = 0$), value of the capacitance density, in units of farads/meter^2, associated with the planar component of the bulk–drain capacitance, while C_{jsw} is the zero bias lineal capacitance, in units of farads/meter, of the aforementioned sidewall areas. The planar drain area, A_d, is

$$A_d = WL_{dif}, \tag{1.151}$$

where L_{dif} is recalled in Figures 1.17 and 1.18 to represent the width of the drain region, which is generally identical to the width of the source implant. Generally, the dimension, L_{dif}, must be extracted empirically from measured data but as a rule of thumb, L_{dif} is nominally of the order of twice the channel length, L. Parameter P_d, is the effective length of the perimeter of the sidewall area and is stipulated by

$$P_d = W + 2L_{dif}. \tag{1.152}$$

Voltage V_j in Equation 1.150 is the built-in potential given by Equation 1.115, while M_j and M_{jsw} are the grading coefficients of the planar and sidewall PN junctions, respectively.* Typically $M_j = 0.5$ and $M_{jsw} = 0.33$. An analogous expression, whose terms convey equally analogous engineering interpretations, prevails for the net bulk–source depletion capacitance, C_{bs}. In particular,

$$C_{bs} = \frac{A_s C_j}{\left(1 - \frac{V_{bs}}{V_j}\right)^{M_j}} + \frac{P_s C_{jsw}}{\left(1 - \frac{V_{bs}}{V_j}\right)^{M_{jsw}}}, \tag{1.153}$$

where in general,

$$\left.\begin{aligned} A_s &\equiv A_d = WL_{dif} \approx 2WL \\ P_s &\equiv P_d = W + 2L_{dif} \approx W + 4L \end{aligned}\right\}. \tag{1.154}$$

1.2.4.2 Gate Capacitances

The second source of MOSFET capacitances is the gate capacitance, which itself is comprised of three distinct components. The first of these components appears between the gate and the bulk substrate. As is apparent from Figure 1.23, this particular capacitance has a very small nonzero frequency value in both weak and strong channel inversion modes, which suggests that the channel inversion layer effectively shields the gate from the bulk substrate. Because the gate–bulk capacitance is invariably very small, it bodes little consequence to achievable MOSFET circuit performance and therefore, it is usually ignored tacitly.

* In HSPICE and other forms of SPICE simulators, the built-in potential, V_j, is symbolized by P_B.

The other two components of net gate capacitance are the gate–source capacitance, C_{gs}, and the gate–drain capacitance, C_{gd}. Each of these energy storage elements is a superposition of an intrinsic module, which derives from the gate, gate oxide, and inverted channel, and an extrinsic constituent, which is attributed to gate oxide overlap at the source and drain sites. Since the inversion layer extends from source-to-drain in only the ohmic regime of operation, different values of these two capacitances prevail for ohmic and saturated operation. The maximum possible intrinsic capacitance established between the gate and the inversion layer is clearly WLC_{ox}. In the ohmic operating regime, this maximum capacitance is partitioned equally between the source and the drain to give identical intrinsic gate–source and gate–drain capacitance values; namely, $WLC_{ox}/2$. Accordingly, in the ohmic regime, the effective gate–source capacitance is

$$C_{gs} = \frac{WLC_{ox}}{2} + WC_{gso}, \tag{1.155}$$

where C_{gso} is the capacitance per unit length associated with the oxide–source overlap. Similarly, the effective gate–drain capacitance in the ohmic operating regime is

$$C_{gd} = \frac{WLC_{ox}}{2} + WC_{gdo}, \tag{1.156}$$

where C_{gdo} is the drain overlap capacitance counterpart to the source overlap region. Typically, C_{gso} and C_{gdo} are as small as 0.25 fF/μm in minimal geometry transistors. Thus, for a transistor characterized by $L = 180$ nm, $W/L = 20$, and an oxide thickness of $T_{ox} = 30$ Å, $C_{gs} = C_{gd} = 4.63$ fF. Observe herewith that the net overlap capacitance is $WC_{gso} = WC_{gdo} = 0.9$ fF, which is almost 20% of the total gate–source (or gate–drain) capacitance.

The capacitance situation in saturation is a bit more intricate than that which prevails in the ohmic regime. In saturated domains where $V_{ds} > V_{dsat}$, pinch off occurs within the source-to-drain channel, thereby leaving an effective depletion zone that is free of mobile carriers near the drain site. Accordingly, the drain–source voltage exerts no influence on the channel charge, and the resultant gate–drain capacitance derives exclusively from the oxide overlap with the drain, that is, the gate–drain capacitance, C_{gd}, in saturation is simply

$$C_{gd} = WC_{gdo}. \tag{1.157}$$

In contrast to the charge depletion prevailing in the channel region adjacent to the drain, a large free carrier population is concentrated near the source. Since this concentration is influenced strongly by interface potential, which is determined by the applied gate–source voltage, it is only logical to expect a comparatively substantial intrinsic gate–source capacitance.

An analytical disclosure of the foregoing gate-to-source capacitance commences with a return to Equations 1.102 and 1.103. If these two equations are combined and if Equation 1.63 is recalled,

$$q[Y_s(\varphi_c(x))][N_s(\varphi_c(x))] = C_{ox}\left[V_{gs} - V_h - \varphi_c(x)\right], \tag{1.158}$$

where the left-hand side of this relationship is understood to be the density of mobile charge in the inversion layer. Upon multiplication of both sides of Equation 1.158 by the gate width, W, the resultant left-hand side of the modified expression represents the net mobile charge per unit length of the inversion layer. It follows that the net differential mobile charge (amassed by electrons in NMOS), say $dq_n[\varphi_c(x)]$, contained in a differential channel volume of depth $Y_s[\varphi_c(x)]$, width W, and length extending from x to $(x + dx)$, is

$$dq_n(\varphi_c(x)) = WC_{ox}\left[V_{gs} - V_h - \varphi_c(x)\right]dx. \tag{1.159}$$

Ignoring mobility degradation incurred by lateral electric fields, Equation 1.104 can be used to recast Equation 1.58 in the form

$$\mathrm{d}q_n(\varphi_c(x)) = \frac{\mu_n(WC_{ox})^2}{I_d}\left[V_{gs} - V_h - \varphi_c(x)\right]^2 \mathrm{d}\varphi_c(x). \tag{1.160}$$

Equation 1.59 can be integrated conveniently from $\varphi_c(0)$ to $\varphi_c(V_{dsat})$, where the indicated interfacial potential limits correspond to the boundaries of the channel inversion layer evidenced in saturation. Such an integration of the left-hand side of Equation 1.59 brackets the net mobile charge, say $Q_n(V_{gs})$, observed in saturation for a stipulated gate–source voltage, V_{gs}. In particular,

$$\begin{aligned} Q_n(V_{gs}) &= \int_0^{V_{dsat}} \mathrm{d}q_n(\varphi_c(x)) = \frac{\mu_n(WC_{ox})^2}{I_d} \int_o^{V_{dsat}} \left[V_{gs} - V_h - \varphi_c(x)\right]^2 \mathrm{d}\varphi_c(x) \\ &= \frac{\mu_n(WC_{ox})^2}{3I_d}\left(V_{gs} - V_h\right)^3, \end{aligned} \tag{1.161}$$

where Equation 1.99 is exploited. Using Equation 1.109 to replace the drain current variable, I_d, in this relationship results in

$$Q_n(V_{gs}) = \frac{2}{3} WLC_{ox}\left(V_{gs} - V_h\right). \tag{1.162}$$

The saturation region value of the intrinsic gate–source capacitance now follows as

$$\frac{\mathrm{d}Q_n(V_{gs})}{\mathrm{d}V_{gs}} = \frac{2}{3} WLC_{ox}, \tag{1.163}$$

whereupon the saturation region value of the net gate–source capacitance, C_{gs}, inclusive of oxide overlap effects at the source site, is

$$C_{gs} = \frac{2}{3} WLC_{ox} + WC_{gso}. \tag{1.164}$$

1.2.4.3 Large-Signal Model

At this juncture, the large signal, or nonlinear, model of an N-channel MOSFET is the structure advanced in Figure 1.34. Depending on whether the transistor undergoing assessment is operated as an ohmic regime or as a saturated device, the equation for the indicated controlled current source, I_d, derives from expressions formulated in Section 1.2.3.2 or Section 1.2.3.3 or, for that matter, the model refinements addressed in Section 1.2.3.4. The depletion capacitances, C_{bs} and C_{bd}, are not affected by the domain of transistor operation, but the appropriate regional values of the capacitances, C_{gs} and C_{gd}, must be culled from the discourse in Section 1.2.4.2.

The model at hand also incorporates four resistive elements. The resistances, r_d and r_s, are respectively associated with the strongly doped drain and source regions, respectively. These resistances are specified in HSPICE simulation software by a sheet resistance parameter, R_{sh}, and drain and source geometric parameters, N_{rd} and N_{rs}. In particular,

$$\left.\begin{aligned} r_d &= N_{rd}R_{sh} \\ r_s &= N_{rs}R_{sh} \end{aligned}\right\}. \tag{1.165}$$

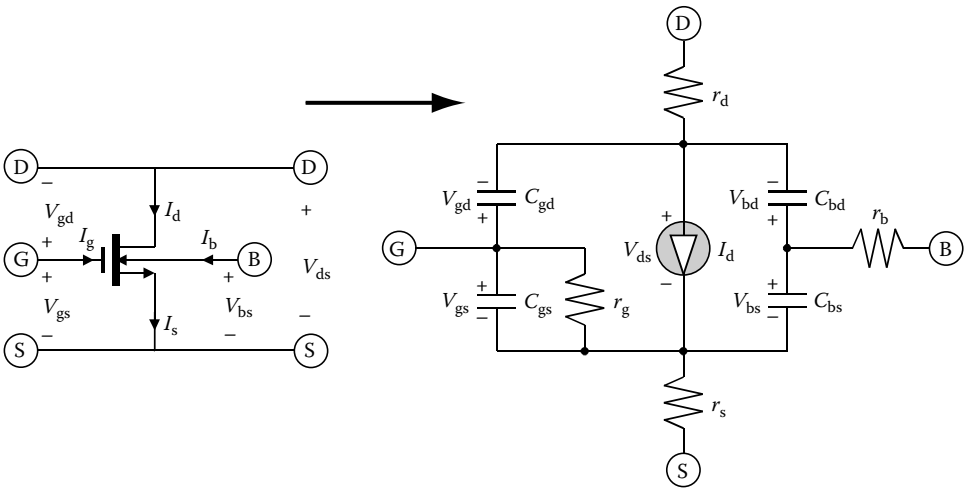


FIGURE 1.34 Large-signal model of an N-channel MOSFET. A topologically identical equivalent circuit prevails for P-channel MOSFETs.

Owing to the high doping concentrations of the drain and the source, which begets a small *sheet resistance* parameter, R_{sh}, resistances r_d and r_s are generally sufficiently small to justify their tacit neglect in most analog circuit applications. In contrast, resistance, r_b, which represents an effective spreading resistance in the bulk substrate, can be as large as the high tens to low hundreds of ohms. Despite its relatively large value, its impact on analog circuit performance is muted by the fact that the bulk rarely conducts significant currents, even at high signal frequencies. However, this resistance does influence the thermal noise characteristics of the drain–source channel.

Like resistance r_b, the gate resistance, r_g, is likewise important from a thermal noise perspective in that it captures the salient effects that thermally agitated mobile charge carriers exert on channel potential. It also looms significant with respect to design problems associated with maximum signal power transfer in radio frequency (RF) circuits [9]. This resistance is computed as [10]

$$r_g = \frac{5}{(\omega C_{gs})^2 R_{ch}}, \tag{1.166}$$

where R_{ch} represents the $V_{ds} = 0$ value of the drain–source channel resistance. Recalling Equation 1.109,

$$R_{ch} = \frac{1}{K_n\left(\frac{W}{L}\right)(V_{gs} - V_h)} = \frac{M_{sat} V_{dsat}}{2I_{dsat}}, \tag{1.167}$$

where Equations 1.129 and 1.132 have been used. Because of the inverse dependence of r_g on the square of radial signal frequency, ω, r_g is infinity for quiescent operating conditions and extremely large for low to even reasonably high frequencies.

1.2.5 Small-Signal Operation

As noted in Section 1.2.1, MOSFETs are the active device of choice in a plethora of high-performance analog integrated circuits. When the fundamental objective of these analog networks is linear I/O signal processing, each MOSFET therein is commonly biased in a saturated regime that ensures, for all applied signals of interest, an instantaneous drain–source voltage, v_{ds}, that is never any smaller than the

instantaneous drain–source saturation voltage, v_{dsat}. To be sure, linear signal processing can also be achieved when transistors operate in their ohmic regimes. But when high performance, in such senses as high gain, wide bandwidth, large dynamic range, and acceptable driving point I/O impedance levels, is a fundamental design objective, saturation is the regime of choice. Accordingly, ohmic linear equivalent circuits of transistors are ignored herewith and left as an investigation exercise for the reader.

A casual inspection of Equation 1.142 suggests that the instantaneous drain current, i_d, flowing in an N-channel MOSFET is a function of three device voltages: the instantaneous gate–source voltage, v_{gs}, the instantaneous drain–source voltage, v_{ds}, and the instantaneous bulk–source voltage, v_{bs} which covertly influences the threshold potential, V_h. An analogous statement applies to P-channel transistors, subject to the current and voltage conventions adopted earlier. Thus, Equation 1.142 can be generalized as

$$i_d \approx \frac{K_n}{2}\left(\frac{W}{L}\right)M_{sat}^2\left(v_{gs}-V_h\right)^2\left(\frac{1+\frac{v_{ds}-v_{dsat}}{V_\lambda}}{1+\frac{v_{gs}-V_h}{V_{ve}}}\right)=f\left(v_{gs},\, v_{ds},\, v_{bs}\right). \tag{1.168}$$

Under zero signal conditions, which is tantamount to operating the considered MOSFET at its quiescent operating point, it is understood that Equation 1.168 yields

$$I_d = f\left(V_{gs}, V_{ds}, V_{bs}\right), \tag{1.169}$$

where the indicated variables in capital letters designate static, or quiescent, device currents and voltages. In other words, the MOSFET described mathematically by Equation 1.169 is in a standby mode that awaits the application of dynamic, invariably time-varying, signals. Prior to signal excitation, the transistor maintains quiescent values of drain current, gate–source voltage, drain–source voltage, and bulk–source voltage that respectively equal I_d, V_{gs}, V_{ds}, and V_{bs}. Signals applied as a current, say I_{ds}, to the drain lead and/or a voltages, say V_1 to the gate–source port, V_2 to the bulk–source port, or V_3 to the drain–source port perturb the quiescent, or *Q-point*, counterparts of these electrical variables to deliver the observable net instantaneous current and voltage responses,

$$\left.\begin{aligned} i_d &= I_d + I_{ds} \\ v_{gs} &= V_{gs} + V_1 \\ v_{bs} &= V_{bs} + V_2 \\ v_{ds} &= V_{ds} + V_3 \end{aligned}\right\}. \tag{1.170}$$

In concert with these relationships, the MOSFET under consideration is said to operate linearly if and only if the signal-induced changes, I_{ds}, V_1, V_2, and V_3, interrelate linearly and if and only if the Q-point currents and voltages are independent of signal strengths. It is crucial to understand that operational linearity in an electronic device does not imply linear relationships among the instantaneous device variables, nor does it imply linearity among the corresponding quiescent values of these variables. Instead, operational linearity implies merely that a selected variable in the selected set of four perturbed variables in Equation 1.170 linearly superimpose with the remaining three electrical signal components.

1.2.5.1 Fundamental Small-Signal Model

Because of the obviously nonlinear nature of Equation 1.168, questions abound as to the plausibility of achieving the aforementioned linearity condition among the electrical perturbations induced by applied signals. Despite its inherently nonlinear nature, Equation 1.168 is a well-behaved functional relationship, which suggests that the desired linearity might be approximated adequately by limiting all signal excursions about respective operating point values to sufficiently small levels. This sufficiently small-signal mandate defines the concept of *small-signal analysis* and produces a *small-signal model* of the

MOSFET. A small-signal analysis reflective of a mathematical exploitation of the corresponding small-signal model are deemed both appropriate and useful if the retention of only the linear terms of the Taylor series expansion of Equation 1.168 about the operating point of the considered device leads to minimal errors in the resultant expression for the signal component of the drain current. Thus,

$$i_d \approx I_d + \left.\frac{\partial i_d}{\partial v_{gs}}\right|_Q (v_{gs} - V_{gs}) + \left.\frac{\partial i_d}{\partial v_{bs}}\right|_Q (v_{bs} - V_{bs}) + \left.\frac{\partial i_d}{\partial v_{ds}}\right|_Q (v_{ds} - V_{ds}), \tag{1.171}$$

where each of the three derivatives on the right-hand side of this relationship are evaluated at the *Q*-point of the MOSFET, that is, at $i_d = I_d$, $v_{gs} = V_{gs}$, $v_{bs} = V_{bs}$, and $v_{ds} = V_{ds}$. Using Equation 1.170 and noting that each of the three subject derivatives is a constant having units of conductance, Equation 1.171 can be couched in the form

$$I_{ds} \approx g_m V_1 + g_{mb} V_2 + \frac{V_3}{r_o}, \tag{1.172}$$

where

$$\left.\begin{aligned} g_m &\triangleq \left.\frac{\partial i_d}{\partial v_{gs}}\right|_Q \\ g_{mb} &\triangleq \left.\frac{\partial i_d}{\partial v_{bs}}\right|_Q \\ \frac{1}{r_o} &\triangleq \left.\frac{\partial i_d}{\partial v_{ds}}\right|_Q \end{aligned}\right\}. \tag{1.173}$$

Equation 1.71 gives rise to the small-signal, low-frequency equivalent circuit depicted in Figure 1.35a. The subject circuit becomes the small-signal, high-frequency MOSFET model if the four capacitances, C_{gs}, C_{gd}, C_{bd}, and C_{bs}, discussed in Section 1.2.4 are appended as indicated in Figure 1.35b. It is important to underscore the fact that either of the models in Figure 1.35 gives no information about the instantaneous electrical variables of a MOSFET, nor does either model allow for the computation of the quiescent values of these variables. Indeed, the models at hand require a priori knowledge of the *Q*-point since the small-signal parameters, g_m, g_{mb}, and r_o, depend on the operating point, as is implied by Equation 1.173. Moreover, the four capacitive elements in the model of Figure 1.35b likewise depend on the *Q*-point at which the considered transistor is biased. In short, the models in Figure 1.35 give first order approximations of the interrelationships among only the small-signal components of the net currents and voltages indigenous to a MOSFET. Although the topologies of both the models drawn in Figure 1.35 pertain to both the ohmic and saturation regimes of N-channel MOSFET operation, the equations to be developed shortly for the low-frequency parameters of these models apply exclusively to the saturation region. Moreover, while Figure 1.35 makes explicit reference to an N-channel transistor, or NMOS, the small-signal, low- and high-frequency equivalent circuits of PMOS units are identical to their NMOS counterparts. This topological identity stems from the fundamental fact that the small-signal models intertwine only signal-induced changes of device currents and voltages about their respective quiescent values. In an attempt to dispel possible confusion, the latter models are offered in Figure 1.36.

In the low-frequency models of either Figure 1.35a or Figure 1.36a the signal component of the bulk current, I_{bs}, flows into an open circuit because the bulk–drain and bulk–source junctions of devices embedded in analog networks are commonly reversed biased. The low-frequency signal component of the gate current, I_{gs}, is very nearly zero because the gate resistance, r_g, is, by Equation 1.166, inversely proportional to the square of the signal frequency. Of course, both of these currents are substantively

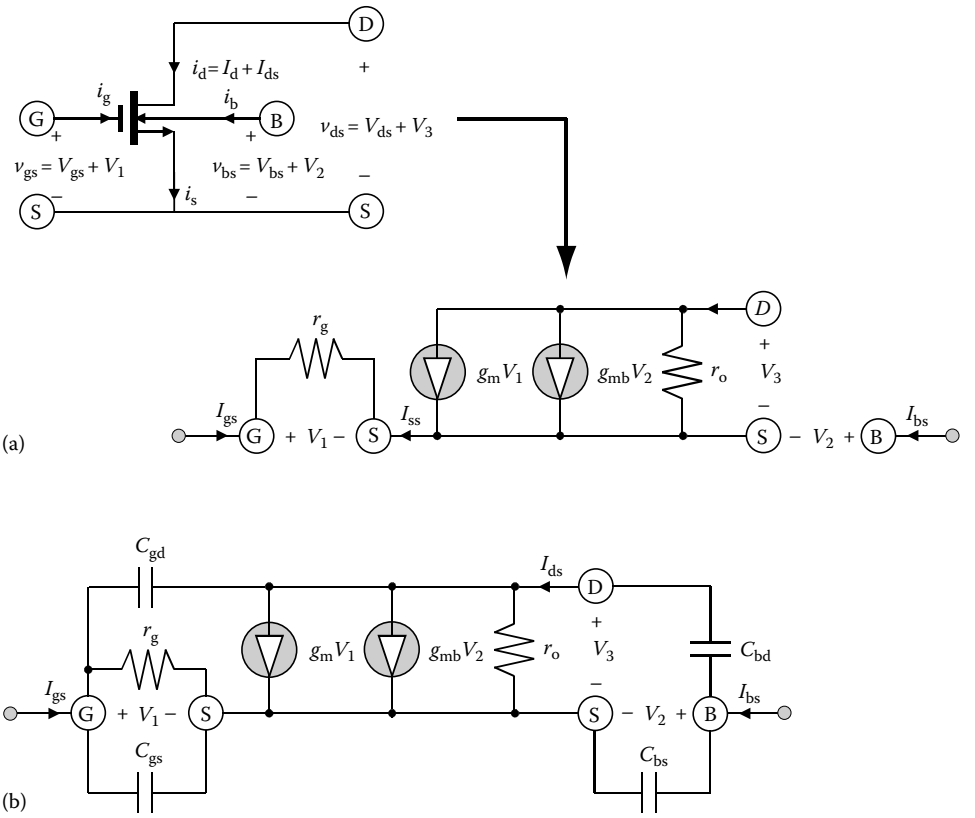


FIGURE 1.35 (a) Small-signal, low-frequency equivalent circuit of an N-channel MOSFET. (b) Small-signal, high-frequency equivalent circuit of an N-channel MOSFET. The topological structures of either model apply to both the ohmic and saturation regimes of operation.

larger at high frequencies where the various capacitances in the models of Figures 1.35b and 1.36b become poor approximations of the open circuits they mirror at low signal frequencies.

The parameter, g_m, is termed the *forward transconductance*. It is a critical analog circuit metric in that it serves as a measure of achievable forward gain. In particular, parameter g_m, when multiplied by the applied gate–source signal voltage, V_1, determines the amount of drain signal current, I_{ds}, manifested by the applied gate–source signal. On the other hand, the *bulk transconductance*, g_{mb}, measures the ability of a MOSFET to transfer applied bulk–source signal, V_2, to the drain signal current response. The controlled current, $g_{mb}V_2$, is negligible when it is much smaller than is its forward transconductance counterpart current, $g_m V_1$. However, it should be noted that depending on the selected quiescent operating point, parameter g_{mb} can be as much as 15%–25% of the forward transconductance, g_m. The current, $g_{mb}V_2$, is entirely inconsequential in numerous analog circuits that configure their utilized MOSFETs in such a way as to operate both the bulk and source terminals at signal ground, which renders $V_2 = 0$. Finally, r_o, the *drain–source channel resistance*, appears as a shunting resistive element across the drain and source terminals. If r_o is infinitely large (which, to be sure, it is not in practical MOSFETs) the drain-to-source small-signal port of a MOSFET behaves as an ideal Norton equivalent current source, that is, the current level determined largely by $g_m V_1$ is unaffected by modulations in the drain–source signal voltage, V_3. It follows that to the extent that the gate–source terminals serve as an input signal port boasting infinitely large impedance and the drain–source terminals function as the output port, the MOSFET emulates an ideal transconductance amplifier if the channel resistance, r_o, is large.

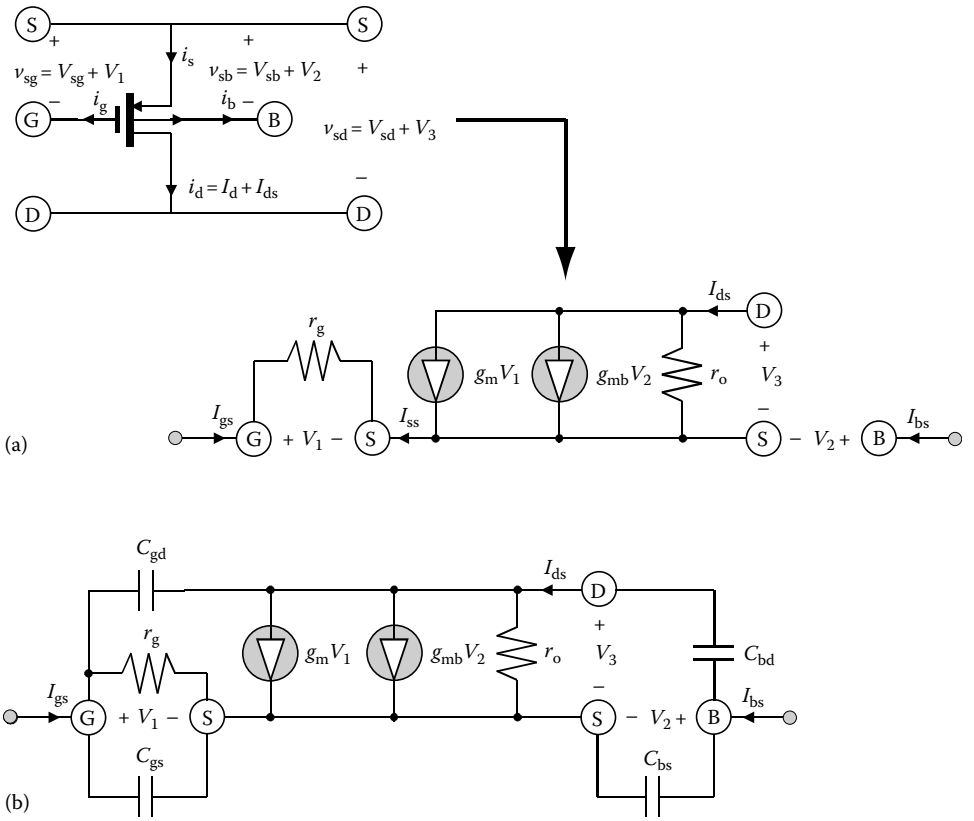


FIGURE 1.36 (a) Small-signal, low-frequency equivalent circuit of a P-channel MOSFET. (b) Small-signal, high-frequency equivalent circuit of a PMOS device.

The determination of the three low-frequency parameters defined in Equation 1.173 requires that the indicated derivatives of the drain current expression in Equation 1.168 be evaluated. This evaluation is an algebraically trying task that borders on a futile engineering enterprise in that many of the physical parameters implicit to Equation 1.168 are rarely disclosed to the circuit designer. It is therefore prudent to condescend to first order approximations of the subject small-signal parameters by replacing Equation 1.168 with the simpler expression,

$$i_d \approx \frac{K_n}{2}\left(\frac{W}{L}\right)(v_{gs} - V_h)^2\left(1 + \frac{v_{ds} - v_{dsat}}{V_\lambda}\right), \tag{1.174}$$

which effectively ignores the influence of both lateral and vertical electric fields in the MOSFET channel. By ignoring the effects of lateral fields, parameter M_{sat} in Equation 1.131 is one, whence the drain saturation voltage in Equation 1.129 is simply the voltage difference, $v_{dsat} = (v_{gs} - V_h)$. Accordingly, Equations 1.173 and 1.174 yield a forward transconductance of

$$g_m \triangleq \left.\frac{\partial i_d}{\partial v_{gs}}\right|_Q \approx \frac{2I_d}{V_{gs} - V_h} - \frac{I_d}{V_\lambda + V_{ds} - V_E}, \tag{1.175}$$

where it is understood that the variables, I_d, V_{gs}, V_h, V_λ, and V_{dsat} reflect the Q-point of the transistor undergoing study. Biasing voltage and standby power constraints ordinarily compel that the transistor be

biased at a drain–source voltage that is only slightly above the drain saturation voltage. Accordingly, V_λ is typically much larger than $(V_{ds} - V_{dsat})$. Moreover, V_λ is generally significantly larger than $V_{dsat}/2$. It follows that the second term on the right-hand side of Equation 1.175 is often negligible, whereupon Equations 1.168 and 1.175 combine for the case of large V_λ to deliver

$$g_m \triangleq \left.\frac{\partial i_d}{\partial v_{gs}}\right|_Q \approx \frac{2I_d}{V_{gs} - V_h} \approx \sqrt{2K_n(W/L)I_d}. \tag{1.176}$$

The result suggests that the forward transconductance of a MOSFET increases with the square root of the product of quiescent drain current and transistor gate aspect ratio. Accordingly, high gain requirements in certain MOSFET amplifiers compel relatively large standby drain currents and/or suitably large gate widths. The former tack conflicts with omnipresent desires for low power operation, while the latter begets increased device capacitances and hence, potentially degraded frequency responses. Observe that while the term in V_λ in Equation 1.175 is usually negligibly small, significant channel length modulation (which translates to small V_λ) is deleterious to high gain objectives.

An evaluation of the bulk transconductance, g_{mb}, requires that Equation 1.96 be considered analytically in conjunction with the threshold voltage term in Equation 1.174. After a bit of messy algebra, it can be shown that

$$g_{mb} \triangleq \left.\frac{\partial i_d}{\partial v_{bs}}\right|_Q \approx \lambda_b g_m, \tag{1.177}$$

where λ_b, which might be termed a *bulk modulation factor*, is

$$\lambda_b = \sqrt{\frac{V_\theta/2}{2V_F - V_{bs}}}. \tag{1.178}$$

Recall a previous assertion to the extent that the bulk transconductance, g_{mb}, may be insignificant in comparison to the small-signal impact of the forward transconductance, g_m. From Equations 1.177 and 1.178, λ_b, and hence g_{mb}, are small if V_θ, the body effect potential defined by Equation 1.90, is small. Since V_θ is proportional to the square of the gate oxide thickness, thin oxide layers conduce small bulk transconductances. It is interesting to note in Equation 1.178 that the small values of the bulk modulation factor that are precipitated by thin oxides are made even smaller by increases in the reverse bias applied between the bulk and source.

The drain–source channel resistance, r_o, in Equation 1.173 is readily confirmed to derive from

$$\frac{1}{r_o} \triangleq \left.\frac{\partial i_d}{\partial v_{ds}}\right|_Q \approx \frac{I_d}{V_\lambda + V_{ds} - V_{dsat}}. \tag{1.179}$$

The result shows that a large channel length modulation voltage, V_λ, gives rise to a large channel resistance, r_o, which in turn implies that the drain–source port of a MOSFET emulates the volt–ampere characteristics of an ideal current source. For conventional values of V_λ, large r_o is seen to require a small drain bias current, I_d.

1.2.5.2 Unity Gain Frequency

The models of Figures 1.35 and 1.36 provide an analytical path for computing a commonly invoked figure of merit for MOSFETs; namely, the *unity gain frequency*, which in radial units is symbolized as ω_T. Although this metric offers a meaningful basis for comparing the high-frequency signal processing

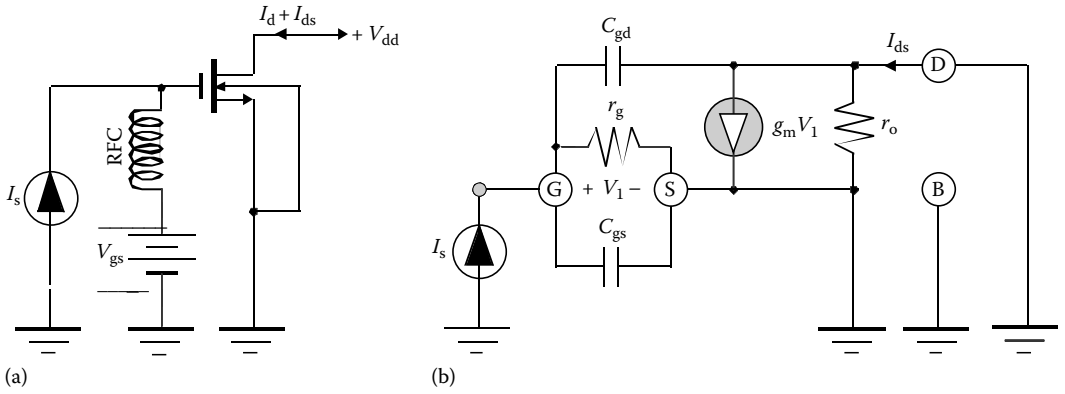


FIGURE 1.37 (a) N-channel common-source MOSFET configured for the evaluation of the unity gain frequency, ω_T, of the transistor. (b) Small-signal, high-frequency equivalent circuit of the network in (a).

capabilities of competing transistors and their associated fabrication processes, its value to bracketing the achievable bandwidths and response speeds of MOSFET circuits is dubious. The latter contention stems from the very definition of the metric. In particular, ω_T is the radial value of signal frequency at which the magnitude of the small-signal, short-circuit current gain of a common-source amplifier degrades to unity. The circuit of relevance is the topology of Figure 1.37a, in which a current signal, I_s is applied to the gate of a MOSFET whose source terminal is incident with signal ground. The radio frequency choke (RFC) provides a conduit for establishing a gate–source bias, V_{gs}, above threshold, while providing a dynamic impedance in series with the gate biasing voltage that is large enough to cajole most of the input signal current to enter the gate terminal.

An input current applied to a gate lead that inherently comprises an open circuit at low frequencies is hardly rational from a circuits perspective. This irrationality is exacerbated by the fact that the drain terminal, where the small-signal current signal response, I_{ds}, to input signal current I_s is extracted, is connected directly to the power supply rail, V_{dd}, thereby rendering the drain terminal short circuited to signal ground (hence the nomenclature, "short circuit" current gain). In other words, the current gain, I_{ds}/I_s, is computed for a common-source amplifier whose gate is driven by signal current and whose drain is short circuited to signal ground, which is hardly a viable analog circuit cell.

Assuming that the transistor at hand operates in saturation, the small-signal equivalent model of the circuit in Figure 1.37a is the structure given in Figure 1.37b. Because the drain, in addition to the bulk and the source, is grounded for signal conditions, the current gain, I_{ds}/I_s, tacitly ignores the high-frequency effects of bulk–drain and bulk–source transistor capacitances. Moreover, the connection of the bulk terminal to the source obviates the need for the bulk transconductance generator, $\lambda_b g_m V_2$, in the model at hand, while short circuiting the drain terminal to the source terminal renders the channel resistance, r_o, inconsequential. Accordingly, an analysis of the structure in Figure 1.37b yields

$$\left.\begin{aligned} I_{ds} &= g_m V_1 - j\omega C_{gd} V_1 \\ I_s &= \frac{V_1}{r_g} + j\omega C_{gs} V_1 + j\omega C_{gd} V_1 \end{aligned}\right\}, \tag{1.180}$$

whence

$$\frac{I_{ds}}{I_s} = \frac{g_m r_g \left(1 - j\omega C_{gd}/g_m\right)}{1 + j\omega r_g \left(C_{gs} + C_{gd}\right)}. \tag{1.181}$$

Since r_g in Equation 1.166 is infinitely large at zero signal frequency, the short circuit gain is seen to be infinity at zero frequency, which reflects engineering intuition in that the gate can conduct no current at zero frequency. In addition, r_g is likely to remain very large in the neighborhood of the 3 dB frequency projected by this gain relationship so that

$$\frac{I_{ds}}{I_s} \approx \frac{g_m}{j\omega\left(C_{gs} + C_{gd}\right)}. \tag{1.182}$$

Equation 1.182 also invokes the reasonable presumption that the frequency, g_m/C_{gd}, of the right half plane zero evidenced on the right-hand side of Equation 1.181 is significantly larger than the aforementioned 3 dB bandwidth. This presumption is tantamount to neglecting the gate-to-drain feedforward through the gate–drain capacitance, C_{gd}, in comparison to the I/O feedforward promoted by the transistor transconductance, g_m. While this approximation is suspect at very high signal frequencies, the approximations leading to Equation 1.182 allow an extrapolated value of the unity gain frequency of

$$\omega_T = 2\pi f_T = \frac{g_m}{C_{gs} + C_{gd}}. \tag{1.183}$$

Clearly, f_T is a highly optimistic estimate of achievable circuit performance, for it pertains expressly to the special case of a drain that is short circuited to the source terminal, thereby quashing the impact on bandwidth of bulk–drain capacitance and any load capacitance that might be driven by the subject transistor. Using Equations 1.106, 1.156, 1.164, 1.176, and 1.174 with V_λ presumed large, the last result is expressible as

$$\omega_T = 2\pi f_T \approx \frac{3\mu_n\left(V_{gs} - V_h\right)}{2L^2\left[1 + \frac{3\left(C_{gso} + C_{gdo}\right)}{2LC_{ox}}\right]}, \tag{1.184}$$

which reflects a direct dependence of the unity gain frequency on free carrier mobility. The result also infers that to the extent that $2LC_{ox} \gg 3(C_{gso} + C_{gdo})$, which may indeed be an engineering stretch for minimal geometry, deep submicron devices, f_T is inversely proportional to the square of channel length. A significant increase in f_T is therefore portended by even a relatively modest shortening of the channel length. The fact that the unity gain frequency value is common corporate banter in the marketing of state of the art transistors and processes arguably underpins the widespread process-engineering penchant for progressively decreased channel lengths.

With the gate resistance, r_g, presumed very large over the signal frequency range of interest, the resultant short circuit current gain in Equation 1.182 is dependent on only frequency-invariant small-signal transistor parameters. Accordingly, Equation 1.183 allows Equation 1.182 to be generalized as the complex frequency domain expression,

$$\frac{I_{ds}}{I_s} \approx \frac{\omega_T}{s}. \tag{1.185}$$

Moreover, the model in Figure 1.35b reduces to the equivalent circuit shown in Figure 1.38a, where signal current I_1 is identified as the current conducted by the gate–source capacitance, C_{gs}. Since signal voltage V_1 is clearly I_1/sC_{gs}, the voltage controlled current, $g_m V_1$, is

$$g_m V_1 = \frac{g_m I_1}{sC_{gs}} = \frac{\omega_T\left(C_{gs} + C_{gd}\right)I_1}{sC_{gs}} = \left(\frac{k_g\omega_T}{s}\right)I_1, \tag{1.186}$$

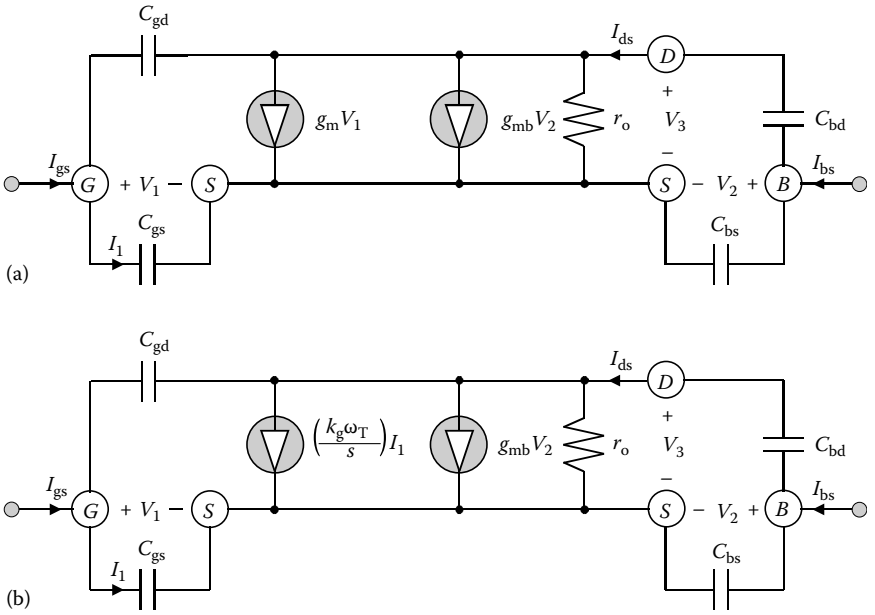


FIGURE 1.38 (a) Small-signal model of Figure 1.35b with gate resistance r_g ignored. (b) Alternative CCCS form of the equivalent circuit in (a).

where

$$k_g = 1 + \frac{C_{gd}}{C_{gs}}. \tag{1.187}$$

Equation 1.186 allows the voltage-controlled current source (VCCS) form of the model shown in Figure 1.38a to be transformed into the equivalent current-controlled structure offered in Figure 1.38b. The latter form proves useful in assessing the performance of amplifiers, such as certain forms of low noise bandpass structures, which utilize source degeneration inductances.

1.2.5.3 Small-Signal Model Development

While the small-signal transistor models shown in Figures 1.35 and 1.36 are topologically correct and conceptually useful from the perspective of linear active network design and first order performance assessment, their engineering utility is limited by two issues. First, the analytical expressions for the parameters embedded in these structures are predicated on a plethora of approximations stemming from the neglect of the effects of lateral and vertical electric fields and simplifications surrounding the charge storage mechanisms of devices and their associated capacitive profiles. These analytically simplifying approximations often place laboratory characterizations of device behavior at odds with physical reality. For example, the substrate doping concentration is not a constant, as is presumed in all foregoing analytical disclosures, but it is less than immediately clear if some sort of weighted average of this dopant level is appropriate for a satisfying voltage–current–charge characterization of a considered transistor. Second, the subject small-signal parameters are dependent on variables, such as carrier mobility, oxide overlap dimensions, doping concentrations, densities of charges trapped in the oxide, regional perimeter dimensions, and the like that are either not released to the circuit designer or are otherwise only vaguely known to the processing foundry.

Because of the foregoing parametric anomalies, reasonably accurate and physically sound assessments of small-signal device and associated circuit performance require that the numerical explication of all relevant small-signal parameters derive from appropriate laboratory measurements conducted on either test device structures or on entire test cells of the circuits undergoing development. A commonly used vehicle toward this characterization end is the *scattering parameters*, or *S-parameters*, measured for a grounded source, grounded gate, or grounded drain interconnection of a subject transistor excited for a suitable range of biasing levels and over an appropriate range of signal frequencies [11]. These parameters are extracted with fixed and known—generally 50 ohm—reference terminations at the input and output ports of the device undergoing test. The measured *S*-parameters, S_{ij}, are then converted into short circuit admittance (*y*-) parameters, y_{ij}. The latter two-port parameters are virtually impossible to discern directly in the laboratory because their numerical delineation mandates the imposition of input and output port signal short circuits, which are difficult to sustain over broad frequency passbands. Once the y_{ij} are determined, it is an involved, but nonetheless straightforward, matter to infer realistic values of most of the parameters implicit to the structures of Figures 1.35 and 1.36.

If the process foundry provides a reliable large-signal HSPICE model, such as the fundamentally heuristic Level 49 MOSFET model, of the device under consideration, the short circuit admittance parameters of the subject transistor can be deduced through appropriate small-signal computer-aided simulations. For example, consider the N-channel MOSFET in Figure 1.39, which is shown connected as a grounded source, three-port configuration. The battery voltage, V_{gg}, biases the gate–source terminals at a greater than threshold value of voltage that establishes the desired quiescent drain current, I_d. Of course, zero quiescent gate current flows in the gate lead of the transistor. On the other hand, the battery voltage, V_{dd}, which modestly influences the quiescent drain current, I_d, is chosen to ensure saturation regime operation of the transistor. Finally, the voltage, V_{bb}, biases the bulk substrate terminal, where it is understood that V_{bb} is ordinarily at most zero. In concert with the traditional stipulations of reverse-biased bulk–source and bulk–drain junctions, zero quiescent gate current is presumed to flow into the bulk. The application of any one or more of the three indicated signal voltages, V_{1s}, V_{2s}, and V_{3s}, produces signal current responses in the gate, bulk, and drain of I_{gs}, I_{bs}, and I_{ds}, respectively. Selecting the "AC" simulation option to manifest a strictly linear HSPICE analysis of the aforementioned current signal responses about respective quiescent values allows the applied signal voltages, V_{1s}, V_{2s}, and V_{3s}, to be set conveniently to amplitudes of 1 V.

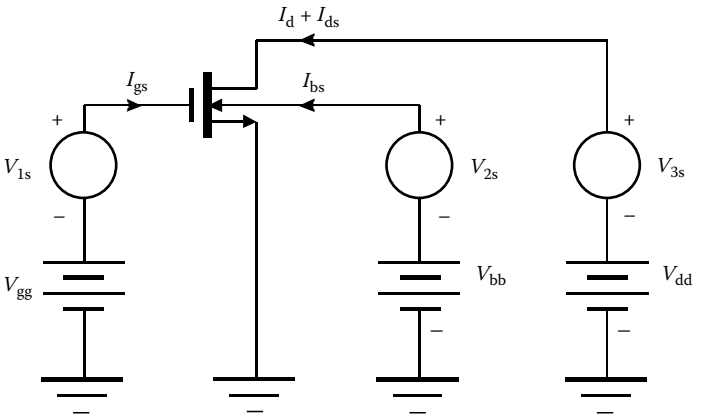


FIGURE 1.39 Common-source test cell of an N-channel transistor configured as a three-port network.

If the applied signal voltages subscribe to the small-signal, linear-operational constraint, superposition theory applies, and the three signal current responses can be described by the linear admittance parameter matrix,

$$\begin{bmatrix} I_{gs} \\ I_{bs} \\ I_{ds} \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} & y_{13} \\ y_{21} & y_{22} & y_{23} \\ y_{31} & y_{32} & y_{33} \end{bmatrix} \begin{bmatrix} V_{1s} \\ V_{2s} \\ V_{3s} \end{bmatrix}. \tag{1.188}$$

In Equation 1.188, the short-circuit admittance parameters, y_{ij}, which are invariably complex numbers, are extracted over signal frequency. In particular,

$$\left.\begin{array}{lll} y_{11} = I_{gs}/V_{1s}|_{V_{2s}=V_{3s}=0}, & y_{12} = I_{gs}/V_{2s}|_{V_{1s}=V_{3s}=0}, & y_{13} = I_{gs}/V_{3s}|_{V_{1s}=V_{2s}=0} \\ y_{21} = I_{bs}/V_{1s}|_{V_{2s}=V_{3s}=0}, & y_{22} = I_{bs}/V_{2s}|_{V_{1s}=V_{3s}=0}, & y_{23} = I_{bs}/V_{3s}|_{V_{1s}=V_{2s}=0} \\ y_{31} = I_{ds}/V_{1s}|_{V_{2s}=V_{3s}=0}, & y_{32} = I_{ds}/V_{2s}|_{V_{1s}=V_{3s}=0}, & y_{33} = I_{ds}/V_{3s}|_{V_{1s}=V_{2s}=0} \end{array}\right\}. \tag{1.189}$$

The real and imaginary parts of all nine of these *y*-parameters can be readily evaluated from a small-signal HSPICE analysis of the structure in Figure 1.39 or alternatively, they can be discerned in terms of scattering parameters gleaned from measurements of a test structure analogous to that of the subject figure.

If the algebraic form of parameter y_{12} in Equation 1.188 is defined as

$$y_{12} \triangleq g_{12} - j\omega C_{12}, \tag{1.190}$$

the first of the equations in Equation 1.188 can be written as

$$I_{gs} = (y_{11} + y_{13} - j\omega C_{12})V_{1s} + g_{12}V_{2s} - y_{13}(V_{1s} - V_{3s}) + j\omega C_{12}(V_{1s} - V_{2s}), \tag{1.191}$$

which can be couched in the form,

$$I_{gs} = \left(\frac{1}{R_i} + j\omega C_i\right)V_{1s} + g_{12}V_{2s} + \left(\frac{1}{R_f} + j\omega C_f\right)(V_{1s} - V_{3s}) + j\omega C_{12}(V_{1s} - V_{2s}). \tag{1.192}$$

The first parenthesized term on the right-hand side of this expression represents a gate-to-ground shunt interconnection of a resistance,

$$R_i = \frac{1}{\mathrm{Re}(y_{11}) + \mathrm{Re}(y_{13})}, \tag{1.193}$$

and a capacitance,

$$C_i = \frac{\mathrm{Im}(y_{11}) + \mathrm{Im}(y_{13})}{\omega} - C_{12}, \tag{1.194}$$

where from Equation 1.190,

$$C_{12} = -\frac{\mathrm{Im}(y_{12})}{\omega}. \tag{1.195}$$

The resistance, R_i, tends to vary as the inverse square of the radial signal frequency. Thus, it is expedient to write R_i as

$$R_i = \frac{K_{Ri}}{\omega^2}, \tag{1.196}$$

where K_{Ri} is a constant boasting the strange dimensions of ohms–(rad/s)^2.

The second term in Equation 1.192 represents a VCCS whose bulk-to-gate transconductance is

$$g_{12} = \mathrm{Re}(y_{12}). \tag{1.197}$$

The second parenthesized factor on the right-hand side of the subject equation connotes a gate-to-drain shunt combination of resistance

$$R_\mathrm{f} = -\frac{1}{\mathrm{Re}(y_{13})} \tag{1.198}$$

and capacitance

$$C_\mathrm{f} = -\frac{\mathrm{Im}(y_{13})}{\omega}. \tag{1.199}$$

As is the case with resistance R_i, R_f also varies as the inverse square of the radial signal frequency. Accordingly,

$$R_\mathrm{f} = \frac{K_\mathrm{Rf}}{\omega^2}. \tag{1.200}$$

Finally, the last term in Equation 1.192 is merely a capacitance, C_{12}, incident between gate and bulk terminals. It is appropriate to interject that over a broad range of signal frequencies that do not exceed the transistor unity gain frequency, f_T, the capacitances, C_i, C_{12}, and C_f, are nearly constants, which suggests that $\mathrm{Im}(y_{11})$, $\mathrm{Im}(y_{12})$, and $\mathrm{Im}(y_{13})$ are nominally linear functions of the radial signal frequency. An analogous statement prevails for all of the other capacitances defined in the forthcoming paragraphs. While resistances R_i and R_f decrease sharply with signal frequency, they are so large (hundreds or even thousands of megohms) that they can usually be neglected in the course of most design-oriented analog circuit analyses.

Letting

$$y_{ij} = g_{ij} - j\omega C_{ij} \tag{1.201}$$

denote the general short circuit admittance parameter, y_{ij}, Equation 1.188 allows the signal current, I_bs, conducted by the bulk to be expressed as

$$I_\mathrm{bs} = (g_{21} - j\omega C_x)V_\mathrm{1s} + \left(\frac{1}{R_\mathrm{bb}} + j\omega C_\mathrm{bb}\right)V_\mathrm{2s} + (g_{23} - j\omega C_{23})V_\mathrm{3s} + j\omega C_{12}(V_\mathrm{2s} - V_\mathrm{1s}), \tag{1.202}$$

where, recalling Equation 1.201,

$$C_x = C_{21} - C_{12} = \frac{\mathrm{Im}(y_{12}) - \mathrm{Im}(y_{21})}{\omega}. \tag{1.203}$$

and

$$g_{21} = \mathrm{Re}(y_{21}). \tag{1.204}$$

The term, $(g_{21} - j\omega C_x)$ in Equation 1.202 is a transadmittance linking the signal gate voltage to the bulk signal current. The second parenthesized term on the right-hand side of Equation 1.202 reflects a bulk-to-ground parallel combination of a frequency-dependent resistance

$$R_{bb} = \frac{1}{\mathrm{Re}(y_{22})} = \frac{K_{Rbb}}{\omega^2}, \tag{1.205}$$

and capacitance

$$C_{bb} = \frac{\mathrm{Im}(y_{22})}{\omega} - C_{12}. \tag{1.206}$$

A second transadmittance factor, $(g_{23} - j\omega C_{23})$, surfaces to model the coupling of the drain signal voltage, V_{3s}, to the bulk signal current, I_{bs}, where

$$\left.\begin{aligned} g_{23} &= \mathrm{Re}(y_{23}) \\ C_{23} &= -\frac{\mathrm{Im}(y_{23})}{\omega} \end{aligned}\right\}. \tag{1.207}$$

Finally, the last term in Equation 1.200 complements its last term counterpart in Equation 1.192 in that it accounts for the bilateral capacitive coupling prevailing between the drain and bulk terminals.

The only current not yet addressed is the signal drain current, I_{ds}. From Equation 1.188,

$$I_{ds} = (g_m - j\omega C_m)V_{1s} + (g_{mb} - j\omega C_{mb})V_{2s} + \left(\frac{1}{R_o} + j\omega C_o\right)V_{3s} + \left(\frac{1}{R_f} + j\omega C_f\right)(V_{3s} - V_{1s}). \tag{1.208}$$

The factor, $(g_m - j\omega C_m)$, is the forward transadmittance that couples the gate signal voltage to the drain signal current. Its constituent variables are

$$\left.\begin{aligned} g_m &= \mathrm{Re}(y_{31}) - \mathrm{Re}(y_{13}) \\ C_m &= \frac{\mathrm{Im}(y_{13}) - \mathrm{Im}(y_{31})}{\omega} \end{aligned}\right\}. \tag{1.209}$$

On the other hand, $(g_{mb} - j\omega C_{mb})$ is the bulk transadmittance serving to bracket the signal drain current response to the signal bulk voltage, V_{2s}. The variables, g_{mb} and C_{mb}, are

$$\left.\begin{aligned} g_{mb} &= \mathrm{Re}(y_{32}) \\ C_{mb} &= -\frac{\mathrm{Im}(y_{32})}{\omega} \end{aligned}\right\}. \tag{1.210}$$

Following Equation 1.177, the bulk modulation factor, λ_b, can be discerned to be

$$\lambda_b = \frac{g_{mb}}{g_m} = \frac{\mathrm{Re}(y_{32})}{\mathrm{Re}(y_{31}) - \mathrm{Re}(y_{13})}. \tag{1.211}$$

The parenthesized factor of the third term on the right-hand side of Equation 1.206 is the drain-to-ground shunt interconnection of resistance R_o and capacitance C_o, such that

$$\left.\begin{aligned} R_o &= \frac{1}{\mathrm{Re}(y_{33}) + \mathrm{Re}(y_{13})} \\ C_o &= \frac{\mathrm{Im}(y_{33}) + \mathrm{Im}(y_{13})}{\omega} \end{aligned}\right\}. \tag{1.212}$$

The last term in Equation 1.208 reflects the previously introduced, bilateral R_f–C_f coupling between the gate and the drain terminals.

Equation 1.192 for the signal gate current, I_{gs}, Equation 1.202 for the signal bulk current, I_{bs}, and Equation 1.208 for the signal drain current, I_{ds}, can now be exploited to develop the foreboding three-port common-source MOSFET model diagrammed in Figure 1.40a. While the model is intractable for manual circuit analysis and considerably more complicated than its simplified brethren in Figures 1.35 and 1.36, it does serve to bolster circuit design insights. First, and perhaps most obviously, the model at hand illustrates the complex interactions of the bulk with the gate, drain, and source regions of a MOSFET. For example, the bulk signal voltage, V_{2s}, precipitates a real controlled source, $\lambda_b g_m V_{2s}$, in addition to a quadrature controlled source, $j\omega C_{mb} V_{2s}$, at the drain–source port. The first of these sources is the expected effect of bulk-induced modulation of MOSFET threshold voltage, but the latter controlled element is slightly south of transparent. The bulk also gives rise to a VCCS, $g_{12} V_{2s}$, in the gate–source port, which accounts for observable bulk-induced increases in high-frequency gate current. These intricacies, together with the complex transadmittance coupling, $(g_{23} + j\omega C_{23}) V_{3s}$, from the drain-to-the bulk seemingly encourage, whenever possible and prudent, operating the MOSFET with its bulk terminal returned to the transistor source terminal. Under such a topological constraint, the model in Figure 1.40a collapses to the almost shockingly simpler network offered in Figure 1.40b.

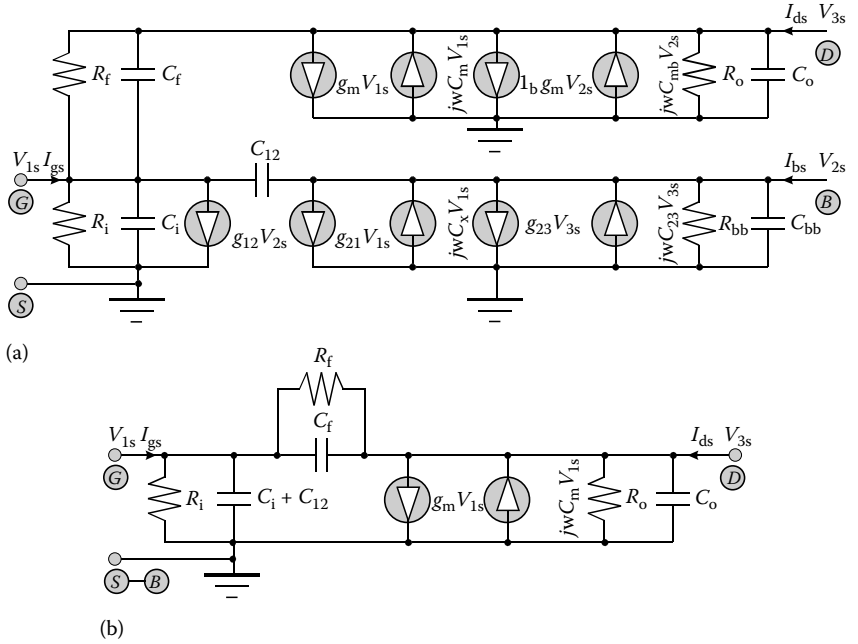


FIGURE 1.40 (a) Small-signal, three-port equivalent circuit for the common-source interconnection of a MOSFET. The three terminal voltages, V_{1s}, V_{2s}, and V_{3s}, denote signal voltages developed with respect to ground at the gate, bulk, and drain terminals, respectively. (b) The equivalent circuit of (a) with the bulk terminal connected directly to the MOSFET source terminal.

A comparison of the model in Figure 1.40b with that of Figure 1.35b suggests that the gate–source resistance, R_i, is effectively the gate resistance, r_g, introduced in Equation 1.166. The resistance, R_f, in Figure 1.40b has no counterpart in Figure 1.35b. Throughout the range of frequencies extending through the unity gain frequency of the considered transistor, both R_i and R_f are so large that they can be ignored for most small-signal analysis ventures, save possibly for a small-signal analysis entailing an assessment of the noise properties of a transistor. The resistance, R_o, is akin to the channel resistance, r_o, in Figure 1.35b. Unlike R_i and R_f, r_o is nominally frequency invariant through the device unity gain frequency metric. With R_i and R_f tacitly ignored and in view of the fact that r_o is independent of frequency, the steady state frequency variable, $j\omega$, in Figure 1.40 can be replaced by the Laplace operator, s, thereby allowing for small-signal step response and other transient investigations of MOSFET amplifiers.

The net capacitance, $(C_i + C_{12})$, in Figure 1.40 is the effective gate–source capacitance, C_{gs}, in Figure 1.35b. Because of the inclusion of capacitance C_{12}, this net gate–source capacitance accounts for gate-to-bulk capacitance, which earlier models presented in this discourse ignore tacitly, primarily because of the high-frequency capacitance characteristics advanced by Figure 1.23. The capacitance, C_f, is the effective gate–drain capacitance, C_{gd}, while capacitance C_o represents the effective bulk–drain capacitance, C_{bd}.

The model in Figure 1.35b highlights a real forward transconductance of g_m, while the models in Figure 1.40 project a complex forward transadmittance, Y_m, of

$$Y_m = g_m - j\omega C_m = g_m e^{j\varphi_m(\omega)} \sqrt{1 + \left(\frac{\omega C_m}{g_m}\right)^2}, \tag{1.213}$$

where

$$\varphi_m(\omega) = -\tan^{-1}\left(\frac{\omega C_m}{g_m}\right) \tag{1.214}$$

denotes an *excess phase angle* associated with the transport of minority carriers in the gate-induced channel extending from the source region-to-the drain region. Equivalently, the angle, $\varphi_m(\omega)$, is associated with an *excess envelope delay*, $T_m(\omega)$, such that

$$T_m(\omega) = -\frac{d\varphi_m(\omega)}{d\omega} = \frac{C_m/g_m}{1 + \left(\frac{\omega C_m}{g_m}\right)^2}, \tag{1.215}$$

whose low-frequency and, in this case, maximum value is obviously (C_m/g_m). Excess delay, for which no account prevails in the simpler models of Figures 1.35 and 1.36, looms potentially critical in feedback circuits in that it acts to degrade the achievable phase margin of the open loop response.

VCCSs having an imaginary transadmittance can be synthesized easily for small-signal, computer-based analyses through the use of a voltage-controlled voltage source (VCVS), a capacitor, and a current-controlled current source (CCCS), as depicted in Figure 1.41a. In this figure, the controlling current, I_i, of the CCCS, αI_i, is

$$I_i = j\omega C\mu V_i, \tag{1.216}$$

whence the indicated controlled current, I_o, is

$$I_o = \alpha I_i = j\omega C\alpha\mu V_i. \tag{1.217}$$

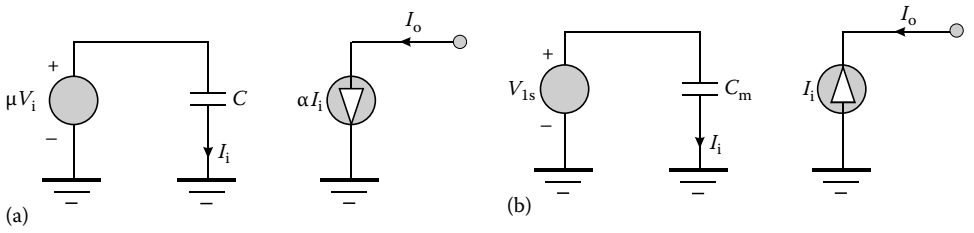


FIGURE 1.41 (a) Synthesis of a VCCS whose transadmittance is imaginary and proportional to radial signal frequency. (b) The synthesis of the controlled current, $-j\omega C_m V_{1s}$, in the models shown in Figure 1.40.

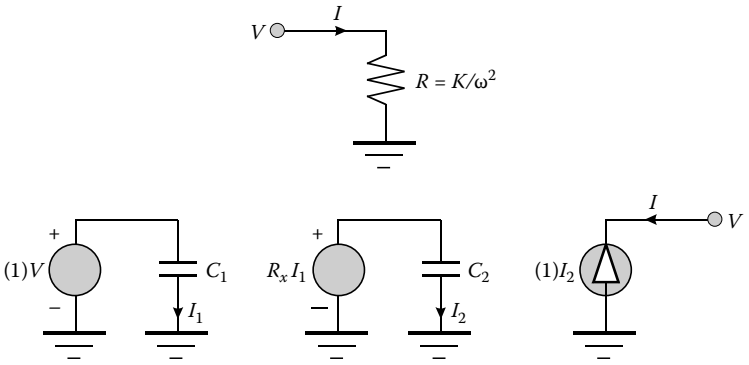


FIGURE 1.42 Synthesis of a branch resistance whose value is inversely proportional to the square of the radial signal frequency. The indicated resistance, R, is synthesized if $R_x = 1/KC_1C_2$.

Thus, for the imaginary component, $j\omega C_m$, of the forward transadmittance, Y_m, in Equation 1.211, $\mu = 1$, $C = C_m$, and $\alpha = -1$ gives the desired controlled current, $-j\omega C_m V_{1s}$, as is abstracted in Figure 1.41b.

Similarly, the frequency variant resistances, R_i, R_f, and R_{bb}, can be synthesized for small-signal, computer-aided analysis purposes using a VCVS, a current-controlled voltage source (CCVS), and a CCCS. This contention is illustrated in Figure 1.42 for the general case of a resistance, R, given by

$$R = -\frac{K}{s^2}, \tag{1.218}$$

which for steady state sinusoidal conditions is the generalized relationship,

$$R = \frac{K}{\omega^2}, \tag{1.219}$$

advanced by Equations 1.196, 1.200, and 1.205. To wit, the controlled current, I_1, generated by the VCVS, $(1)V$, is $I_1 = sC_1V$, while the current, I_2, established in response to the CCVS, (R_xI_1), is $I_2 = sC_2R_xI_1 = s^2C_1C_2R_xV$. It follows that the resistance, R, presented to the port driven by the CCCS, $(1)I_2$, is

$$R = \frac{V}{-I_2} = \frac{V}{-s^2C_1C_2R_xV} = -\frac{1}{s^2C_1C_2R_x}. \tag{1.220}$$

For arbitrary values of capacitances C_1 and C_2, selecting

$$R_x = \frac{1}{KC_1C_2} \tag{1.221}$$

achieves the desired resistance value set forth by Equation 1.218.

Parameterization example:
An N-channel transistor featuring a channel length of 180 nm has the Level 49 HSPICE parameters that appear in Table 1.2. The transistor is implemented with a gate aspect ratio of $W/L = 25$, and is biased at $V_{gs} = 1.1$ V, $V_{ds} = 1$ V and $V_{bs} = 0$ V. The device undergoing study is earmarked for analog small-signal applications that embrace a signal frequency range extending from 100 MHz to 10 GHz. For this frequency passband, determine nominal values of all of the parameters indigenous to the small-signal, common-source model of Figure 1.40b. Also, compute the extrapolated unity gain frequency of the transistor at the given quiescent operating point. Express these results as maximum value, minimum value, average value, and standard deviation (referred to the average value) over a frequency passband extending from 100 MHz to 10 GHz.

TABLE 1.2 Representative Level 49 HSPICE Parameters for an NMOS Transistor in a Fabrication Process Featuring a Nominal Channel Length of 180 nM

Model 180 nM NMOS (Level = 49)

+VERSION = 3.1	TNOM = 27	TOX = 4E-9	XJ = 1E-7
+NCH = 2.3549E17	VTH0 = 0.3627858	K1 = 0.5873035	K2 = 4.793052E-3
+K3 = 1E-3	K3B = 2.2736112	W0 = 1E-7	NLX = 1.675684E-7
+DVT0W = 0	DVT1W = 0	DVT2W = 0	DVT0 = 1.7838401
+DVT1 = 0.5354277	DVT2 = −1.243646E-3	U0 = 263.3294995	UA = −1.359749E-9
+UB = 2.250116E-18	UC = 5.204485E-11	VSAT = 1.083427E5	A0 = 2
+AGS = 0.4289385	B0 = −6.378671E-9	B1 = −1E-7	KETA = −0.0127717
+A1 = 5.347644E-4	A2 = 0.8370202	RDSW = 150	PRWG = 0.5
+PRWB = −0.2	WR = 1	WINT = 1.798714E-9	LINT = 7.631769E-9
+XL = −2E-8	XW = −1E-8	DWG = −3.268901E-9	DWB = 7.685893E-9
+VOFF = −0.0882278	NFACTOR = 2.5	CIT = 0	CDSC = 2.4E-4
+CDSCD = 0	CDSCB = 0	ETA0 = 2.455162E-3	ETAB = 1
+DSUB = 0.0173531	PCLM = 0.7303352	PDIBLC1 = 0.2246297	PDIBLC2 = 2.220529E-3
+PDIBLCB = −0.1	DROUT = 0.7685422	PSCBE1 = 8.697563E9	PSCBE2 = 5E-10
+PVAG = 0	DELTA = 0.01	RSH = 6.7	MOBMOD = 1
+PRT = 0	UTE = −1.5	KT1 = −0.11	KT1L = 0
+KT2 = 0.022	UA1 = 4.31E-9	UB1 = −7.61E-18	UC1 = −5.6E-11
+AT = 3.3E4	WL = 0	WLN = 1	WW = 0
+WWN = 1	WWL = 0	LL = 0	LLN = 1
+LW = 0	LWN = 1	LWL = 0	CAPMOD = 2
+XPART = 0.5	CGDO = 716E-12	CGSO = 716E-12	CGBO = 1E-12
+CJ = 9.725711E-4	PB = 0.7300537	MJ = 0.365507	CJSW = 2.604808E-10
+PBSW = 0.4	MJSW = 0.1	CJSWG = 3.3E-10	PBSWG = 0.4
+MJSWG = 0.1	CF = 0	PVTH0 = 4.289276E-4	PRDSW = −4.2003751
+PK2 = −4.920718E-4	WKETA = 6.938214E-4	LKETA = −0.0118628	PU0 = 24.2772783
+PUA = 9.138642E-11	PUB = 0	PVSAT = 1.680804E3	PETA0 = 2.44792E-6
+PKETA = 4.537962E-5)			

Results:

1. Before proceeding with the simulation, the planar source and drain areas, A_s and A_d, as well as the source and drain peripheral dimensions, P_s and P_d, must be computed through an appeal to Equation 1.154. These are $A_s = A_d = (1.62)(10^{-12})$ m^2 and $P_s = P_d = (5.22)(10^{-6})$ *m*. In arriving at these figures, use is made of the fact that for a channel length of $L = 180$ nm and a gate aspect ratio of $W/L = 25$, the gate width is $W = 4.5$ μm. The parameters, L, W, A_s, A_d, P_s, and P_d are inserted directly on the model line of the HSPICE net list. For example, the model line used in the simulations executed herewith is

$$M2\ 6\ 4\ 0\ 7\ 180\ \text{nM}, \quad L = 180\text{n}, \quad W = 4.5\ \text{u}, \quad A_S = 1.62\ \text{p}, \quad A_D = 1.62\ \text{P},$$
$$P_S = 5.22\ \text{u}, \quad P_D = 5.22\ \text{u}.$$

   In this model line, *M*2 identifies the transistor undergoing examination, ''6'' is the number of the drain node, ''4'' is the gate node number, ''0'' is the number of the grounded source node, and ''7'' is the number of the bulk substrate node. The insert, ''180 nM,'' identifies the name of the model used for the subject transistor.

   An HSPICE simulation of the simple test cell shown in Figure 1.39 can now be straightforwardly executed. In this test structure, $V_{gg} = 1.1$ V, $V_{dd} = 1$ V, and $V_{bb} = 0$ V combine to set the desired operating point of the transistor. The operating point information disclosed by the static HSPICE simulation is as follows.

   ID 8.9407E-04 (drain current is $I_d = 894.1$ μA)
   IS -8.9407E-04 (source current flows out of device and is virtually identical to the drain current)
   IB -1.0002E-12 (bulk current is about one picoampere and flows out of the device)
   IBD -9.9417E-13 (bulk current is sum of the bulk–drain and bulk–source junction currents)
   IBS -6.0237E-15 (bulk current is sum of the bulk–drain and bulk–source junction currents)
   VGS 1.1000 (desired gate–source quiescent voltage)
   VDS 1.0000 (desired drain–source quiescent voltage)
   VBS 0.0000 (desired bulk–source quiescent voltage)
   VTH 0.5102 (simulated threshold voltage is $V_h = 510.2$ mV)
   VDSAT 0.3149 (simulated drain saturation voltage is $V_{dsat} = 314.9$ mV)

   It should be noted that the quiescent drain source voltage, $V_{ds} = 1.0$ V, is certainly larger than the simulated drain saturation voltage, $V_{dsat} = 314.9$ mV. Accordingly, the device at hand operates in its saturation regime for suitable small-signal excitations.

2. A small-signal, computer-aided simulation of the test circuit in Figure 1.39 can now be executed at the quiescent operating point established in the preceding step of this exercise. The objective of this simulation is to ascertain the real and imaginary components of each of the nine short circuit admittance parameters, y_{ij}, introduced in Equation 1.188. The model parameters then derive from the pertinent equations given in Section 1.2.5.3.

   Gate-to-drain resistance coefficient, K_{Rf}:
   Maximum value is $(7.93)(10^{27})$
   Minimum value is $(7.92)(10^{27})$
   Average value is $(7.93)(10^{27})$
   Standard deviation is 0.03%

   Gate-to-drain capacitance, C_f:
   Maximum value is 3.22 fF
   Minimum value is 3.21 fF

Average value is 3.22 fF
Standard deviation is 0.02%

Gate-to-source resistance coefficient, K_{Ri}:
Maximum value is $(4.25)(10^{27})$
Minimum value is $(4.24)(10^{27})$
Average value is $(4.25)(10^{27})$
Standard deviation is 0.03%

Gate-to-source capacitance, C_i:
Maximum value is 7.56 fF
Minimum value is 7.55 fF
Average value is 7.55 fF
Standard deviation is 0.02%

Bulk–gate transconductance, g_{12}:
Maximum value is 1.01 μmho
Minimum value is 0 μmho
Average value is 0.11 μmho
Standard deviation is 191.97%

Gate-to-bulk capacitance, C_{12}:
Maximum value is 0.42 fF
Minimum value is 0.42 fF
Average value is 0.42 fF
Standard deviation is 0%

Forward transconductance, g_m:
Maximum value is 2.03 mmho
Minimum value is 2.03 mmho
Average value is 2.03 mmho
Standard deviation is 0.01%

Transadmittance capacitance, C_m:
Maximum value is 2.00 fF
Minimum value is 1.99 fF
Average value is 2.00 fF
Standard deviation is 0.05%

Gate–bulk transconductance, g_{21}:
Maximum value is 0.82 μmho
Minimum value is 0 μmho
Average value is 0.18 μmho
Standard deviation is 191.97%

Gate–bulk transadmittance capacitance, C_x:
Maximum value is 0.68 fF
Minimum value is 0.67 fF
Average value is 0.67 fF
Standard deviation is 0.03%

Bulk transconductance modulation factor, λ_b:
Maximum value is 0.21
Minimum value is 0.21

Average value is 0.21
Standard deviation is 0.03%

Drain–bulk transconductance, g_{23}:
Maximum value is 0 μmho
Minimum value is −0.41 μmho
Average value is −0.05 μmho
Standard deviation is 191.97%

Bulk transadmittance capacitance, C_{mb}:
Maximum value is 3.19 fF
Minimum value is 3.18 fF
Average value is 3.19 fF
Standard deviation is 0.02%

Drain–bulk transadmittance capacitance, C_{23}:
Maximum value is 2.63 fF
Minimum value is 2.63 fF
Average value is 2.63 fF
Standard deviation is 0.02%

Drain–source channel resistance, R_o:
Maximum value is 10.38 kΩ
Minimum value is 10.33 kΩ
Average value is 10.37 kΩ
Standard deviation is 0.09%

Bulk-to-source resistance coefficient, K_{Rbb}:
Maximum value is $(7.03)(10^{27})$
Minimum value is $(6.78)(10^{27})$
Average value is $(6.99)(10^{27})$
Standard deviation is 0.76%

Drain–source capacitance, C_o:
Maximum value is 2.62 fF
Minimum value is 2.62 fF
Average value is 2.62 fF
Standard deviation is 0.03%

Bulk–source capacitance, C_{bb}:
Maximum value is 6.83 fF
Minimum value is 6.82 fF
Average value is 6.83 fF
Standard deviation is 0.02%

3. Equation 1.183 is the pertinent equation for the computation of the extrapolated unity gain frequency. To this end, the average forward transconductance has been computed to be $g_m = 2.03$ mmho. The effective gate–source capacitance, C_{gs}, is the computed average value, $C_i = 7.55$ fF, which accounts for gate–source overlap and any other second order phenomena embraced by the utilized HSPICE model. On the other hand, the effective average gate–drain capacitance, C_{gd}, is $C_f = 3.22$ fF, which, like C_i, incorporates all pertinent high order device characterization phenomena. Accordingly

$$f_T = \frac{g_m}{2\pi(C_i + C_f)} = 30.0 \text{ GHz}.$$

Comments: With the exception of parameters g_{12}, g_{21}, and g_{23}, the quoted standard deviation numbers indicate an excellent model fit to circuit theoretic issues. These three transconductances can also be made to agree well with theoretical disclosures if they are each allowed to vary as the square of the radial signal frequency. However, their values are so small as to render overt concern of them unproductive.

The computed unity gain frequency, f_T, is within range of the expected frequency performance of representative MOSFETs manufactured in a 180 nM technology process. It is interesting to note, however, that the effective gate–drain capacitance (3.22 fF), which is traditionally ignored in first order, high-frequency circuit analysis ventures, is, in this case, almost 43% of the effective gate–source capacitance (7.55 fF).

1.2.6 Design-Oriented Analysis Strategy

When a MOSFET is exploited for a linear analog signal processing application, an essential early design requirement entails the implementation of suitable biasing. Generally, this biasing must ensure that for all pertinent signal levels, each transistor used to supply gain, impedance conversion, constant current, constant voltage, or other I/O properties operates in its saturated domain where its drain–source voltages, V_{ds}, is at least as large as its drain saturation voltage, V_{dsat}. When $V_{ds} \geq V_{dsat}$, Equation 1.142 is the applicable relationship for ascertaining a gate–source voltage, V_{gs}, commensurate with a target drain current, I_d, conducted at a given or desired value of drain–source voltage.

Unfortunately, academic satisfaction does not often resonate with the engineering reality that underlies predictable, reliable, and reproducible integrated circuit design. For the biasing issue at hand, Equation 1.142 is fraught with numerous shortfalls. Despite its algebraic cumbersomeness, Equation 1.142 is only an approximation of the static volt–ampere characteristics of a MOSFET operated in saturation, owing to a variety of analytical liberties exploited with respect to charge storage, charge transport, carrier mobility, and the other phenomenological issues discussed in preceding sections. Even if Equation 1.142 were an accurate disclosure of the aforementioned static characteristics, challenges surround its utilization because circuit and system designers are rarely privy to the physical and process parameters on which the metrics, K_n, V_h, V_{dsat}, V_{ve}, V_{le}, and V_λ, are dependent. These model variables can be discerned reliably through only laboratory measurement of static device responses or via analyses conducted on appropriate computer-based simulations founded on accurate and reliable transistor models.

On the tacit presumption that the foregoing six model variables can be extracted satisfactorily from measurement and/or simulation, Equation 1.142 might be supplanted by the more familiar, nominally square law relationship,

$$I_d = \frac{K_{ne}}{2}\left(\frac{W}{L}\right)\left(V_{gs} - V_h\right)^2\left(1 + \frac{V_{ds} - V_{dsat}}{V_\lambda}\right), \tag{1.222}$$

where K_{ne} symbolizes the effective transconductance coefficient,

$$K_{ne} = \frac{K_n M_{sat}^2}{1 + \frac{V_{gs} - V_h}{V_{ve}}} \approx \frac{K_n}{\left(1 + \frac{V_{gs} - V_h}{V_{ve}}\right)\left[1 + 0.78\left(\frac{V_{gs} - V_h}{V_{le}}\right)\right]}. \tag{1.223}$$

This effective transconductance coefficient accounts for mobility degradation deriving from strong vertical (gate-to-channel) electric fields through the variable, V_{ve}, as well as mobility degradation caused by lateral (drain-to-source) electric fields, which is monitored by variable V_{le}. While Equation 1.222 suggests a relatively straightforward square law dependence of drain current on the so called *excess,* or *effective, gate–source voltage*, $(V_{gs} - V_h)$, particularly for the commonly encountered situation of $(V_{ds} - V_{dsat}) \ll V_\lambda$, it should be noted that K_{ne} is inversely proportional to a quadratic function of the

excess gate–source voltage. Typically, V_{ve} is of the order of 5- to 20-fold the value of V_{le} and thus, the possibility of simplifying Equation 1.223 to ease computational strain, while preserving computational accuracy, is dubious.

Example:

An N-channel transistor featuring a channel length of 180 nM has the Level 49 HSPICE parameters given in Table 1.2. The transistor is to be biased in saturation at $V_{ds} = 1$ V and $I_d \approx 1$ mA to achieve a small-signal transconductance, g_m, of at least 3 mmhos. Assuming that the bulk terminal is incident with the transistor source terminal, choose a reasonable gate aspect ratio, W/L, determine the required gate–source voltage bias, V_{gs}, and estimate the model parameters implicit to Equation 1.222.

Results:

1. The applicable circuit for computer-aided investigation is offered in Figure 1.43, where the transistor model parameters are those that appear in Table 1.2, and the gate aspect ratio, W/L, is to be determined. The null voltage source in the drain circuit of the device facilitates the extraction of the quiescent drain current, I_d. It is understood that for biasing purposes, the area and perimeter parameters, A_s, A_d, P_s, and P_d, are of no consequence and can therefore be defaulted to any convenient value. Initially, set $V_{gs} = 1$ V and $W/L = 1$ and, of course, $V_{ds} = 1$ V. The HSPICE static simulation reveals $I_d = 46.4$ μA, $V_{dsat} = 262.8$ mV, $V_h = 519.7$ mV, and $g_m = 136.5$ μmho. Since $V_{gs} = 1$ V is certainly larger than $V_h = 519.7$ mV and $V_{ds} = 1$ V $> V_{dsat} = 262.8$ mV, the transistor is clearly turned on and operates in its saturation domain.
2. With $W/L = 1$, the simulated drain current is a factor of 21.55 times smaller than the target current of 1 mA. This observation seemingly suggests the need for increasing the gate aspect ratio from 1 to 21.55, since the drain current is ostensibly proportional to W/L. In truth, the actual drain current is not directly proportional to W/L because of numerous second order effects, including weak dependencies of threshold voltage, drain saturation voltage, and parameter M_{sat} on gate width W. Experience shows that a more viable gate aspect ratio adjustment is about twice that computed or in this case, about 40. With $W/L = 40$ and $V_{gs} = V_{ds} = 1$ V, HSPICE delivers $I_d = 1.08$ mA, $V_{dsat} = 278.2$ mV, $V_h = 510.2$ mV, and $g_m = 3.17$ mmho. The simulated transconductance value satisfies its design target. Although V_{gs} can be decreased modestly to reduce the drain current to 1 mA, this exercise is unnecessary in view of the effects of routinely encountered device processing

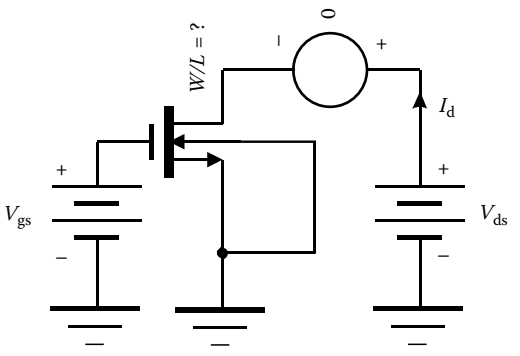


FIGURE 1.43 Circuit structure for MOSFET biasing simulation. The Level 49 HSPICE parameters of the transistor are delineated in Table 1.2.

vagaries and model parameter uncertainties. Thus the design requirement is satisfied for $W/L = 40$ and $V_{gs} = V_{ds} = 1$ V.

3. The model parameterization exercise begins by using Equation 1.141 to compute the voltage, V_{ve}. From Table 1.2, the oxide thickness is $T_{ox} = 4(10^{-9})$ m, which is 40 Å. Accordingly, $V_{ve} = 40/15 = 2.667$ V.
4. The next step in the parameterization process entails operating the transistor undergoing study at a V_{ds} value that equals its saturated value of 278.2 mV. This tack reduces the last parenthesized factor on the right-hand side of Equation 1.222 to unity, thereby simplifying the computation of the effective transconductance parameter, K_{ne}. With $W/L = 40$, $V_{gs} = 1$ V, and $V_{ds} = V_{dsat} = 278.2$ mV, HSPICE produces $I_d = 878.33$ μA and $V_h = 510.0$ mV. Appealing to Equation 1.222, parameter K_{ne} follows forthwith as $K_{ne} = 182.9$ μmho/V.
5. Recalling that $V_{dsat} = 278.2$ mV and $(V_{gs} - V_h) = (1 - 0.510)$ V $= 0.490$ V, Equation 1.129 delivers $M_{sat} = 0.5678$. The previously documented approximate equation, Equation 1.136, relating M_{sat} to variable α can be used to determine the numerical value of α for $V_{gs} = 1$ V and $V_h = 510.0$ mV. Alternatively, Equation 1.131 can be solved for α directly to yield

$$\alpha = \frac{2(1 - M_{sat})}{M_{sat}^2} = 2.682. \tag{E1.1}$$

   Using Equation 1.130, parameter V_{le} follows forthwith as $V_{le} = 182.7$ mV.
6. With $K_{ne} = 182.9$ μmho/V, $V_{ve} = 2.667$ V, $M_{sat} = 0.5678$, $V_{gs} = 1$ V, and $V_h = 510.0$ mV, the device transconductance parameter, K_n, follows from Equation 1.223 as $K_n = 671.7$ μmho/V. It is interesting to observe that the effective transconductance factor, K_{ne}, is almost 3.7 times smaller that the "actual" transconductance coefficient, K_n. Experience testifies to the apparent fact that for deep submicron devices, $2.5 \leq K_n/K_{ne} \leq 4$ is typical.
7. In principle, V_{ve}, V_{le}, V_h, V_{dsat}, K_n, and thus K_{ne}, do not vary with changes in the drain–source voltage, V_{ds}. Accordingly, the ratio of the drain current (1.08 mA) for $V_{ds} = 1$ volt to the drain current (878.33 μA) at $V_{ds} = V_{dsat} = 278.2$ mV is solely attributed to the last parenthesized factor on the right-hand side of Equation 1.222, that is,

$$\frac{I_d|_{V_{ds}=1\text{ V}}}{I_d|_{V_{ds}=V_{dsat}}} = \frac{1.08\text{ mA}}{878.33\ \mu\text{A}} = 1.230 = 1 + \frac{V_{ds} - V_{dsat}}{V_\lambda}. \tag{E1.2}$$

   It follows that the channel length modulation voltage is $V_\lambda = 3.144$ V.
8. In an attempt to demonstrate the propriety of the foregoing modeling exercise, the forward static transfer characteristic of the subject transistor is modeled in HSPICE for both $V_{ds} = 1$ V and $V_{ds} = 1.5$ V. The simulated results are then compared with calculations deriving from Equations 1.222 and 1.223 using the computed values of V_{ve}, V_{le}, and V_λ and the simulated disclosures for W/L, V_{dsat}, and V_h. Specifically, $V_{ve} = 2.667$ volts, $V_{le} = 182.7$ mV, $V_\lambda = 3.144$ volts, $W/L = 40$, $V_{dsat} = 278.2$ mV, and $V_h = 510.0$ mV.

Figures 1.44 and 1.45 display the results of the foregoing comparative study. In Figure 1.44, the simulated and calculated forward transistor characteristics in the saturation domain are displayed for a drain–source voltage, V_{ds}, of 1.0 V. The calculations corroborate reasonably well with pertinent simulations in that ±15% error is observed for $0.73\text{ V} < V_{gs} < 1.89$ V. It is notable that $V_{gs} = 0.73$ V is only slightly larger than 200 mV above threshold level, while at $V_{gs} = 1.89$ V, the transistor no longer operates in its saturation domain when $V_{ds} = 1$ V. Figure 1.45 confirms better corroboration between calculated and simulated results for $V_{ds} = 1.5$ V. In particular, the computational error is within ±9% for $0.92\text{ V} < V_{gs} < 2$ V and is within ±15% for $0.78\text{ V} < V_{gs} < 2$ V.

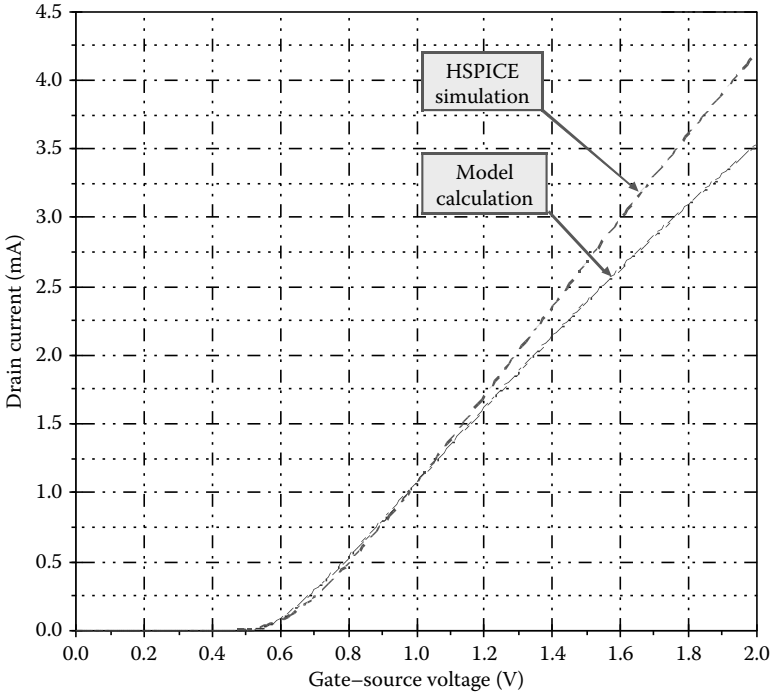


FIGURE 1.44 Simulated and calculated forward static transfer characteristic for the NMOS transistor whose model parameters are delineated in Table 1.2. The transistor is operated at a drain–source voltage, V_{ds}, of 1 V.

Comments: In Step #2 of the foregoing computational procedure, the gate aspect ratio, W/L, is the pivotal metric for achieving the desired transconductance and transistor drain current. If power dissipation is a dominant design concern, W/L can be increased above the value of 40 discerned in this example, with the understanding that the gate–source voltage, V_{gs}, can be reduced commensurately, thereby reducing the static drain current and hence, the power dissipation of the transistor. Of course, the primary penalty of large gate aspect ratio is a possible degradation of high-frequency circuit response since, as is confirmed by Equation 1.154, the capacitance area and peripheral dimensions increase in proportion to the gate width, W.

In Step #3, the metric, V_{ve}, is evaluated in terms of a purely empirical, and indeed crude first order, relationship to the oxide thickness, T_{ox}. A possible way around this dilemma is to compute V_{ve} and all of the other requisite modeling parameters by curve fitting Equation 1.222 to simulated or actually measured static data. While this approach may be academically satisfying, it may be imprudent from a design time perspective. Keep in mind that biasing is not the fundamental performance objective of an analog circuit; rather, biasing is the necessary condition that expedites the desired analog responses.

The drain saturation voltage, V_{dsat}, is obviously a nonlinear function of the excess gate voltage, $(V_{gs}-V_h)$, owing to the parameter, M_{sat}. But in addition, V_{dsat} changes slightly with the applied drain–source voltage, V_{ds}. Indeed, the Level 49 model parameters account for a slight sensitivity of threshold voltage on V_{ds}, which is as anticipated since the interface potential throughout the entire channel varies somewhat as a function of the lateral (drain-to-source) field engendered by V_{ds}.

Finally, it should be noted that the computed value (3.144 V) of the channel length modulation voltage, V_λ, is appreciably smaller than values often propounded in the textbook literature. However, V_λ is indeed a relatively small voltage for deep submicron MOS technology transistors. This anemic voltage is the principle cause of correspondingly small drain–source channel resistances, which renders the

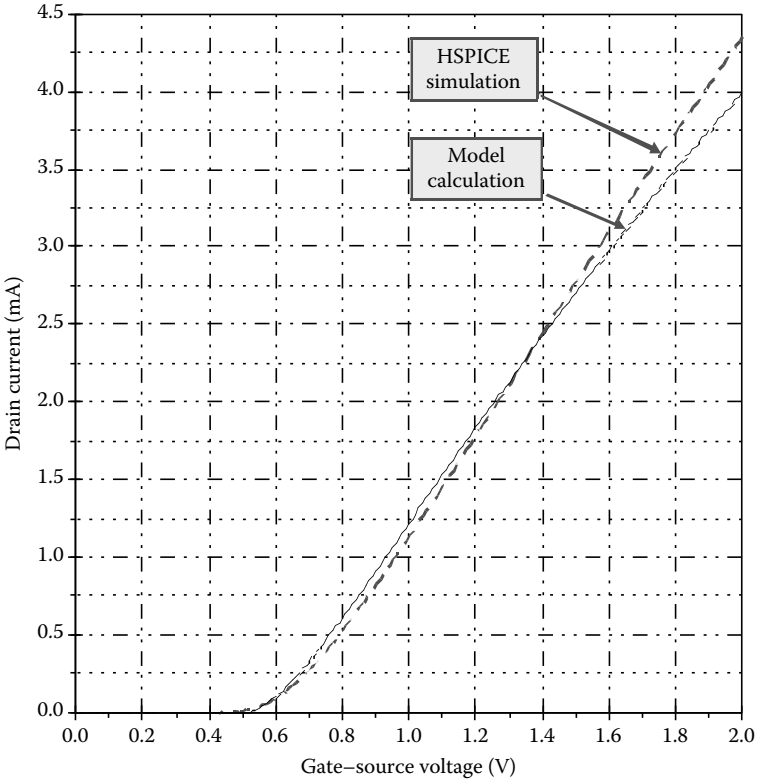


FIGURE 1.45 Simulated and calculated forward static transfer characteristic for the NMOS transistor whose model parameters are delineated in Table 1.2. The transistor is operated at a drain–source voltage, V_{ds}, of 1.5 V.

realization of transconductor amplifiers, as might be used in operational transconductor amplifier-capacitor (OTA-C) filters, a daunting challenge. The desire for accuracy surrounding the enumeration of V_λ is exacerbated by the fact that parameter V_λ is not the constant that is presumed tacitly in the foregoing demonstration. Instead, and as is suggested by Equation 1.114, V_λ is functionally dependent on drain–source voltage, drain saturation voltage, and threshold voltage. If V_λ or the drain–source channel resistance is critical in an analog circuit design endeavor, care must therefore be exercised to ensure that model parameters are extracted in terms of measured or simulated data that largely mirror the desired or expected operating state of the utilized transistor.

References

1. S. W. Sze, *Physics of Semiconductor Devices*. New York: John Wiley & Sons, 1969, pp. 366–379.
2. A. B. Glaser and G. E. Subak-Sharpe, *Integrated Circuit Engineering: Design, Fabrication, and Applications*. Reading, MA: Addison-Wesley Publishing Company, 1977, pp. 80–94.
3. A. S. Grove, *Physics and Technology of Semiconductor Devices*. New York: John Wiley & Sons, 1967, pp. 263–285.
4. A. Goetzberger, Ideal MOS curves for silicon, *Bell System Technical Journal*, 45, 1097, 1966.
5. S. R. Hofstein and G. Warfield, Physical limitation on the frequency response of a semiconductor surface inversion layer, *Solid State Electronics*, 8, 321, 1965.
6. A. S. Grove, E. H. Snow, B. E. Deal, and C. T. Sah, Simple physical model for the space-charge capacitance of metal–oxide–semiconductor structures, *Journal of Applied Physics*, 33, 2458, 1964.

7. R. L. Geiger, P. E. Allen, and N. R. Strader, *VLSI Techniques for Analog and Digital Circuits*. New York: McGraw-Hill Publishing Company, 1990, pp. 174–177.
8. D. Johns and K. Martin, *Analog Integrated Circuit Design*. New York: John Wiley & Sons, Inc, 1997, pp. 24–27.
9. T. H. Lee, *The Design of CMOS Radio-Frequency Integrated Circuits*, 2nd Ed. Cambridge, United Kingdom: Cambridge University Press, 2004, Chaps. 11 and 12.
10. A. van der Ziel, *Noise in Solid State Devices and Circuits*. New York: John Wiley & Sons, Inc, 1986.
11. J. Choma and W.-K. Chen, *Feedback Networks: Theory and Applications*. Singapore: World Scientific Press, 2007, Chap. 3.

1.3 JFET, MESFET, and HEMT Technology and Devices

Stephen I. Long

1.3.1 Introduction

Many types of field effect devices are used in analog IC and RFIC design. Section 1.2 described the MOSFET and associated device models. MOSFETs are currently the predominant field effect device used in analog circuit applications due to the pervasive CMOS technology. CMOS fabrication is relatively inexpensive when not scaled below 0.25 μm. However, mask costs for 130 nm and below increase very rapidly, limiting applications to only those requiring extremely high volume. Also, drain breakdown voltage is quite low, of the order of 1 V for 65 nm CMOS. This constrains dynamic range or power output in certain applications.

Other field effect devices are available, but are considered niche market devices in most cases. This would include the legacy silicon JFET technology, still used in conjunction with bipolar transistors for some lower frequency analog applications. Compound semiconductor-based field effect devices (MESFET, HEMT, p-HEMT, m-HEMT) are often the FET of choice for applications requiring very wide bandwidth, extremely low noise, high gain at mm-wave frequencies, and high output powers at frequencies above 2 GHz. Cost of fabrication is frequently less than that of CMOS in smaller volume applications because the mask set costs are typically an order of magnitude less. Also, the compound semiconductor devices are grown on semi-insulating substrates. Passive components such as spiral inductors, MIM capacitors and deposited resistors have less parasitic capacitance and higher Q than is typical for silicon-based RFICs.

In this section, the silicon JFET and the main compound semiconductor HEMT devices will be described. Special emphasis will be placed on the GaN HEMTs whose performance is exceptionally good for microwave and mm-wave power amplifiers.

1.3.2 Silicon JFET Device Operation and Technology

Although the silicon JFET is today a legacy device, it is still used in some bipolar analog ICs to provide an inexpensive BiFET IC technology. Also, the description of its current–voltage characteristic is similar to any FET which uses a pn or Schottky metal–semiconductor junction for the gate electrode. The JFET consists of a conductive channel with source and drain contacts whose conductance is controlled by a gate electrode. The channel can be fabricated in either conductivity type, n or p, and both normally-on (depletion mode) and normally-off (enhancement mode) type devices are possible. The circuit symbols typically used for JFETs are shown in Figure 1.46 along with the bias polarities of active region operation for these four device possibilities. For analog circuit applications, the depletion mode is almost exclusively utilized because it provides a larger range of input voltage and therefore greater dynamic range. In silicon,

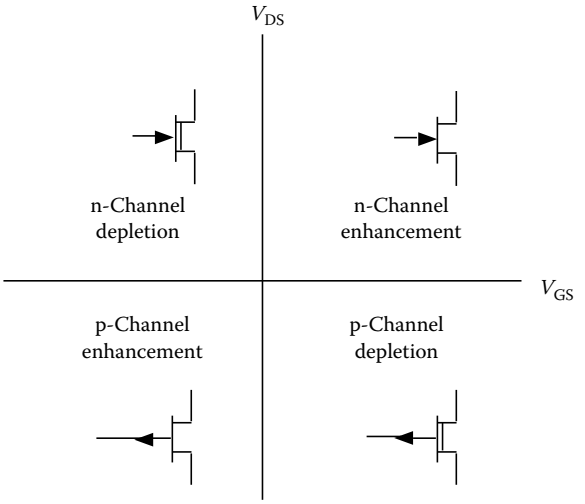


FIGURE 1.46 The circuit symbols typically used for JFETs are shown with the bias polarities for active region operation.

both p- and n-channel JFETs are used, but when compound semiconductor materials such as GaAs or InGaAs are used to build the FET, n-channel devices are used almost exclusively.

When fabricated with silicon, the JFET is used in analog IC processes for its high input impedance, limited by the depletion capacitance and leakage current of a reverse-biased pn junction. When the JFETs are used at the input stage, an op-amp with low input bias current, at least at room temperature, can be built. Fortunately, a p-channel JFET can be fabricated with a standard bipolar process with few additional process steps. This enables inexpensive BiFET processes to be employed for such applications. Unfortunately, the simple process modifications required for integrating JFETs and BJTs are not consistent with the requirements for high-performance devices. Short-gate lengths and high-channel doping levels are generally not possible. So the transconductance per channel width and the gain–bandwidth product of JFETs integrated with a traditional analog BJT process are not very good. The short-circuit current gain–bandwidth product (f_T) is about 50 MHz for an integrated p-channel JFET. The MOSFETs in a BiCMOS process are much better devices, however, a BiCMOS process does not often include both NPN and PNP BJTs needed for high-performance analog circuits.

Discrete silicon JFETs are available with much better performance because they can be fabricated with a process optimized for the JFET. Typical applications are for low-noise amplifiers up to the VHF/UHF range. Noise figures less than 0.1 dB can be obtained at low frequencies with high source impedances and 2 dB at high frequencies at the noise matching input condition with high performance discrete silicon JFETs. The low input gate current, I_G, which can be in the picoamp range, causes the shot noise (proportional to $\sqrt{I_G}$) component to be very low. The input equivalent noise current of the JFET is mainly due to input referred channel (Johnson) noise. This property gives very low noise performance when presented with a high source impedance. In this case, the JFET is often superior to a BJT for noise. For low source impedances, the BJT is generally better.

Compound semiconductor materials such as GaAs and InGaAs are used to fabricate JFET-like devices called metal-semiconductor FET (MESFETs) and high electron mobility transistor (HEMTs). The reason for using these materials is superior performance at high frequencies. These devices are unequaled for gain–bandwidth, ultralow noise, and power amplification at frequencies above 10 GHz and up to 300 GHz. Integrated analog microwave circuits are fabricated with these devices and are commercially available for use in low noise receiver and power amplifier applications. Some representative results will be summarized in Table 1.5.

1.3.2.1 JFET Static *I–V* Characteristics

The JFET differs in structure and in the details of its operation from the MOSFET discussed in Section 1.3.2. Figure 1.47 shows an idealized cross section of a JFET. The channel consists of a doped region, which can be either p- or n-type, with source and drain contacts at each end. The channel is generally isolated from its surrounding substrate material by a reverse biased p–n junction. The depletion regions are bounded in Figure 1.47 by dashed lines and are unshaded. The thin, doped channel region forms a resistor of width W into the page and height d. A gate electrode is located at the center of the channel, defined by a semiconductor region of opposite conductivity type of length L. An n-channel structure is shown here for purposes of illustration. The p-type gate constricts the channel, both through the depth of the diffusion or implant used to produce the gate and through the depletion layer formed at the p–n junction. The height of the channel can be varied by biasing the gate relative to the source (V_{GS}). A reverse bias increases the depletion layer thickness, reducing the channel height and the drain current. If V_{GS} is large enough that the channel is completely depleted, the drain current will become very small. This condition corresponds to the cutoff and subthreshold current regions of operation, and the V_{GS} required to cut-off the channel is called V_P, the pinch-off voltage. V_P corresponds to the threshold voltage that was defined for the MOSFET. Similarly, a forward bias between gate and channel can be used to increase drain current, up to the point where the gate junction begins to conduct. Most JFETs are designed to be depletion-mode (normally on); drain current can flow when $V_{GS} = 0$ and they are normally operated with a reverse-biased gate junction. It is also possible, however, to fabricate enhancement-mode JFETs by use of a thinner or more lightly doped channel.

The pinch-off voltage is a sensitive function of the doping and thickness of the channel region. It can be found if the channel-doping profile, $N(x)$, is known through Poisson's equation. For a nonuniform profile,

$$V_P = V_{BI} - \frac{q}{\varepsilon}\int_0^d xN(x)\mathrm{d}x \tag{1.224}$$

For uniform doping, $N(x) = N_D$ and the familiar result in Equation 1.225 shows that the pinch-off voltage depends on the square of the thickness. This result shows that very precise control of profile depth is needed if good matching and reproducibility of pinch-off voltage is to be obtained [7].

$$V_P - V_{BI} - \frac{qN_D d^2}{2\varepsilon} \tag{1.225}$$

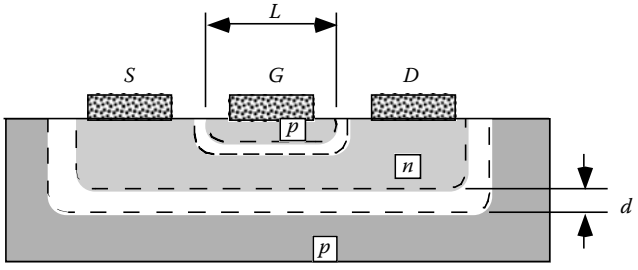


FIGURE 1.47 Idealized cross section of a JFET. The depletion regions are bounded with dashed lines and are unshaded.

1.3.2.2 JFET Operating Regions

The static current–voltage characteristics of the JFET can be categorized by the five regions of operation shown in Figure 1.48 for an n-channel device. The mechanisms that produce these regions can be qualitatively understood by referring to the channel cross sections in Figure 1.49. In these figures, the doped channel region is shaded, and the depletion region is white. First, consider the JFET in Figure 1.49a with small V_{DS} ($\ll V_{GS} - V_P$). This condition corresponds to the ohmic region (sometimes called linear or triode region) where current and voltage are linearly related. At small drain voltages, the depletion layer height is nearly uniform, the electric fields in the channel are too small to saturate the carrier velocity, and thus the channel behaves like a linear resistor. The resistance can be varied by changing V_{GS}. The channel height is reduced by increasing the reverse bias on the gate leading to an increased resistance.

As V_{DS} increases, the depletion layer thickness grows down the length of the channel as shown in Figure 1.49b. This occurs because the drain current causes a voltage increase along the channel as it flows through the channel resistance. Since the depletion layer thickness is governed by the gate-to-channel voltage (V_{GC}), there is an increasing reverse bias that leads to constriction of the channel at the drain end of the gate. Ideally, when $V_{DS} = V_{GS} - V_P$, then $V_{GC} = V_P$, and the channel height will approach zero (pinch-off). The constricted channel will cause the drain current to saturate as shown. Further increases in V_{DS} do not cause the drain current to increase since the channel has already constricted to a minimum height and the additional potential is accommodated by lateral extension of the depletion region at the drain end of the gate. This region of operation is generally described as the pinch-off region (rather than the saturation region in order to avoid confusion with BJT saturation). The height of the channel is not actually zero but is limited by the mobile channel charge, which travels at saturated drift velocity in this high field region.

If $V_{GS} < 0$, then the initial channel height at the source is reduced, I_D is less, and the pinch-off region occurs at a smaller drain voltage $V_{DS} = V_{GS} - V_P$. The saturation of drain current can also occur at smaller V_{DS} if the gate length is very small. In this case, the electric field in the channel is large, and the carrier velocity will saturate before the channel can reach pinch-off. Velocity saturation will also limit drain current.

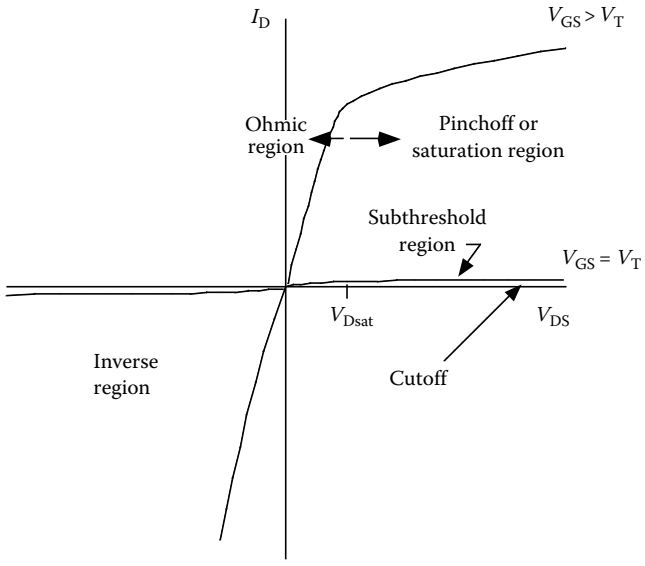


FIGURE 1.48 The static current–voltage characteristics of the JFET can be categorized by five regions of operation. An n-channel device is shown in this illustration.

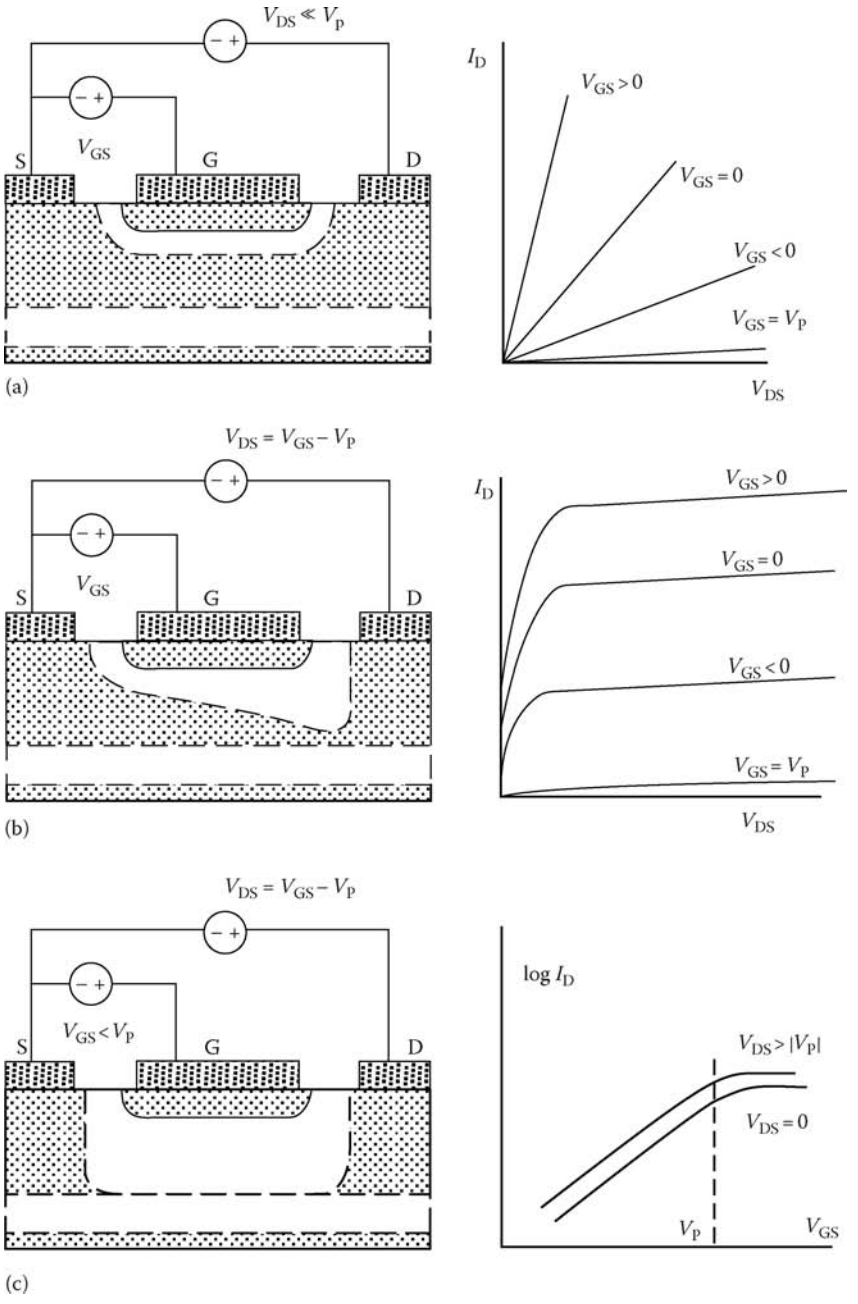


FIGURE 1.49 (a) Ohmic region with small V_{DS} ($\ll V_{GS} - V_P$). (b) When $V_{DS} = V_{GS} - V_P$, the channel height will become narrow at the drain end of the gate. The device enters pinch-off. The constricted channel will cause the drain current to saturate as shown. (c) Cutoff and subthreshold current regions occur when the depletion region extends through the channel.

The subthreshold region of operation, shown in Figure 1.49c is defined when small drain currents continue to flow even though $V_{GS} \leq V_P$. While technically this gate bias should produce cutoff, some small fraction of the electrons from the source region will have sufficient energy to overcome the

potential barrier caused by the gate depletion region and will drift into the drain region and produce a current. Since the energy distribution is exponential with potential, the current flow in this region varies exponentially with V_{GS}.

The inverse region occurs when the polarity of the drain bias is reversed. This region is of little interest for the JFET since gate-to-drain conduction of the gate diode limits the operation to the linear region only.

1.3.2.3 Channel-Length Modulation Effect

A close look at the *I–V* characteristic in the pinch-off region shows that the incremental conductivity or slope of this region is not equal to zero. There is some finite slope that is not expected from the simple velocity saturation or pinch-off models. Channel length modulation is one explanation for this increase; the position under the gate where pinch-off or velocity-saturation first occurs moves toward the source as V_{DS} increases. This is due to the expansion of the drain side depletion region at large V_{DS}. Figure 1.50 illustrates this point. Here, a channel cross section is shown for $V_{DS} = V_{GS} - V_P$ in Figure 1.50a and for $V_{DS} \gg V_{GS} - V_P$ in Figure 1.50b. While pinch-off always occurs when the gate-to-channel voltage is V_P, the higher drain voltage causes the location of this point ($x = L$) to move closer to the source end of the channel. Since the electric field in this region, E, is roughly proportional to $(V_{GS} - V_P)/L$ where L is now a function of V_{DS} and V_{GS} and the carrier velocity $v = \mu E$ (by assumption), then the current must increase as the channel length decreases due to increasing carrier velocity. If the channel length is short, velocity saturation may cause the drain current to saturate. In this case, the velocity saturation point moves closer to the source as drain voltage is increased. Since the length has decreased, less gate-to-channel voltage is needed to produce the critical field for velocity saturation. Less voltage implies a wider channel opening, hence more current.

1.3.2.4 Temperature Effects

There are two mechanisms that influence the drain current of the JFET when temperature is changed [8,9]. First, the pinch-off voltage becomes more negative (for n-channel) with increase in temperature, therefore requiring lower V_{GS} to cut off the channel or to enter the pinch-off region. Therefore, when the device is operating in the pinch-off region, and $V_{GS} - V_P$ is small, the drain current will increase with temperature. This effect is caused by the decrease in the built-in voltage of the gate-to-channel junction with increasing temperature. Second, the carrier mobility and saturated drift velocity decreases with temperature. This causes a reduction in drain current that is in opposition to the first effect. This effect dominates for large $V_{GS} - V_P$. Therefore, there is a V_{GS} value for which the drain current is exactly compensated by the two effects. This is illustrated qualitatively in Figure 1.51.

The gate current is also affected by temperature, as it is the reverse current of a pn junction. The current increases roughly by a factor of 2 for each 10°C increase in temperature. At high temperatures, the input current of a JFET input stage may become comparable to that of a well-designed BJT input stage of an op-amp, thus losing some of the benefit of the mixed BJT–JFET circuit design.

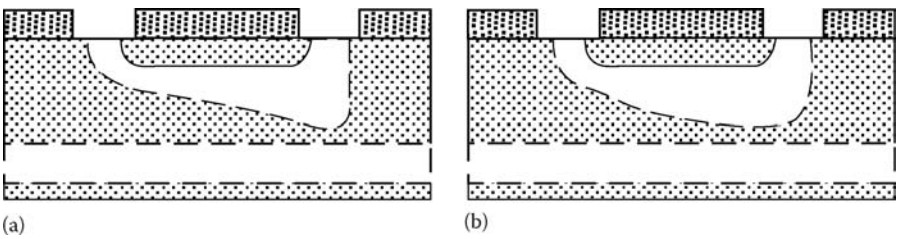


FIGURE 1.50 A channel cross section is shown for $V_{DS} = V_{GS} - V_P$ in (a) and for $V_{DS} \gg V_{GS} - V_P$ in (b). While pinch-off always occurs when the gate-to-channel voltage is V_P, the higher drain voltage causes the location of this point ($x = L$) to move closer to the source end of the channel.

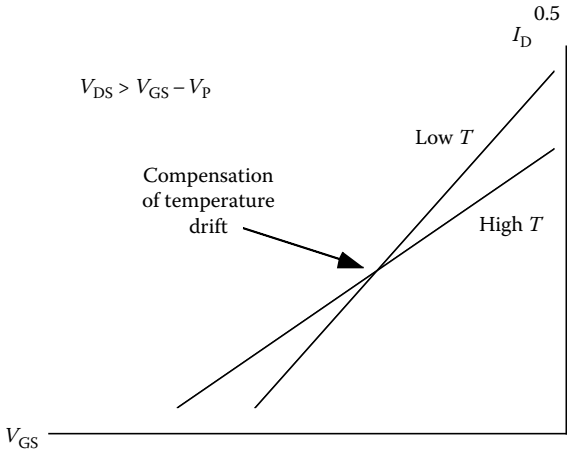


FIGURE 1.51 Effect of temperature on the drain current in the pinch-off region.

1.3.2.5 JFET Models

Most applications of the JFET in analog ICs employ the pinch-off region of operation. It is this region that provides power gain and buffer (source follower) capability for the device, so the models for the JFET presented below will concentrate on this region. It will also be assumed that the gate–source junction will not be biased into forward conduction. Although forward conduction is simple to model using the ideal diode equation within the FET equivalent circuit models, this bias condition is not useful for the principal analog circuit applications of the JFET and will also be avoided in the discussion that follows.

1.3.2.5.1 Large-Signal Model: Drain Current Equations

Equations modeling the large signal JFET $I_D - V_{GS}$ characteristic can be derived for the two extreme cases of FET operation in the pinch-off region. A gradually decreasing channel height and mobility limited drift velocity in the channel are appropriate assumptions for very long gate length FETs. A fixed channel height at pinch-off with velocity saturation limited drift velocity are more suitable for short gate lengths.

The square-law transfer characteristic [10] given by Equation 1.226 provides a good

$$I_D = I_{DSS}\left(1 - \frac{V_{GS}}{V_P}\right)^2 (1 + \lambda V_{DS}) \tag{1.226}$$

approximation to measured device characteristics in the case of long gate length (>5 μm) or very low electric fields in the channel $(V_{GS} - V_P)/L < E_{sat}$. In both cases, the channel height varies slowly and the velocity remains proportional to mobility. E_{sat} is the critical field for saturation of drift velocity, about 3.5 kV/cm for GaAs and 20 kV/cm for Si. I_{DSS} is defined as the drain current in the pinch-off region when $V_{GS} = 0$. The first two terms of the equation are useful for approximate calculation of DC biasing. The third term models the finite drain conductance caused by the channel length modulation effect. The parameter λ in this term is derived from the intercept of the drain current when extrapolated back to zero as shown in Figure 1.52.

Equation 1.226 is also used to represent the pinch-off region in the SPICE JFET model. It is parameterized in a slightly different form as shown below in Equation 1.227.

$$I_D = \beta(V_{GS,i} - V_{T0})^2 (1 + \lambda V_{DS}) \tag{1.227}$$

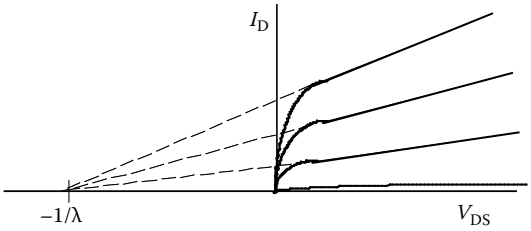


FIGURE 1.52 The channel length modulation parameter λ is defined by the extrapolation of the drain current in saturation to $I_D = 0$.

These equations are the same if $V_{T0} = V_P$, and

$$\beta = \frac{I_{DSS}}{V_P^2} \tag{1.228}$$

and

$$\begin{aligned} V_{GS,i} &= V_{GS} - I_D R_S \\ V_{DS,i} &= V_{DS} - I_E(R_S + R_D) \end{aligned} \tag{1.229}$$

The pinch-off region is defined for $V_{DS,i} \geq V_{GS,i} - V_{T0}$ as is usual for the gradual channel approximation. R_S and R_D are the parasitic source and drain resistances associated with the contacts and the part of the channel that is outside of the gate junction. These resistances will reduce the internal device voltages below the applied terminal voltages as shown in Equations 1.229.

For shorter gate length devices, improved models have been proposed and implemented in SPICE3 and some of the many commercial SPICE products, often in the MESFET model. The Statz model [11] is frequently used for this purpose. This model modifies the drain current dependence on V_{GS} by adding a velocity saturation model parameter b in the denominator as shown in Equation 1.230.

$$I_D = \left[\frac{\beta(V_{GS,i} - V_{T0})^2}{1 + b(V_{GS,i} - V_{T0})}\right](1 + \lambda V_{DS,i}) \tag{1.230}$$

This added term allows the drain current to be nearly square law in V_{GS} for small $V_{GS} - V_{T0}$, but it becomes almost linear when V_{GS} is large, effectively emulating the rapid rise in transconductance followed by saturation that is typical in short channel devices. Although the specific behavior of the drain current is sensitive to the vertical doping profile in the channel), Equation 1.230 is flexible enough to accommodate most short channel device characteristics with uniform or nonuniform channel doping. Another feature of short gate length FETs that this model predicts adequately is a saturation of I_D at $V_{DS,i} < V_{GS,i} - V_{T0}$. This early transition into the pinch-off region is also a consequence of velocity saturation and is widely observed.

1.3.2.5.2 Small-Signal Model

The small-signal model for the JFET in the pinch-off region is shown in Figure 1.53. The voltage dependent current source models the transconductance g_m as a constant which can be derived from the drain current equations above from

$$g_m = \frac{\partial I_D}{\partial V_{GS}} \tag{1.231}$$

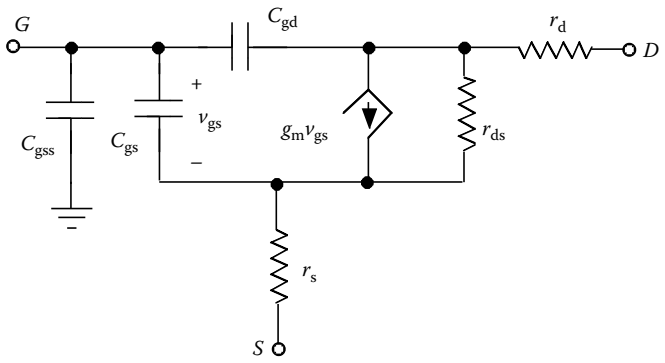


FIGURE 1.53 Small-signal model for the JFET in the pinch-off region.

The square-law current model (Equation 1.226) predicts a linearly increasing g_m with V_{GS}

$$g_m = -\frac{2I_{DSS}}{V_P}\left(1 - \frac{V_{GS}}{V_P}\right) \tag{1.232}$$

whereas a model which includes some velocity saturation effects such as Equation 1.230 would predict a saturation in g_m.

The small-signal output resistance, r_o, models the channel length modulation effect. This is also derived from the drain current equations through

$$r_o^{-1} = \frac{\partial I_D}{\partial V_{DS}} \tag{1.233}$$

For both models, r_o is determined by

$$r_o = \frac{1}{I_D \lambda} \tag{1.234}$$

The small-signal capacitors representing the nonlinear, voltage-dependent C_{gs}, C_{gd}, and C_{gss} are also shown in Figure 1.53. Parasitic source and drain resistances, R_S and R_D can also be included, as shown. If they are not included in the small-signal model, the effect of these parasitics can sometimes be produced in the intrinsic FET model by reducing the intrinsic g_m of the device.

The short-circuit current gain–bandwidth product, f_T, defined in Equation 1.235 is a high-frequency figure of merit for transistors. It is inversely proportional to the transit time τ of the channel charge, and it is increased by reducing the gate length. Reduced L also reduces the gate capacitance and increases transconductance. The material also affects f_T as higher drift velocity leads to higher g_m.

$$f_T = \frac{g_m}{2\pi\left(C_{gss} + C_{gs} + C_{gd}\right)} = \frac{1}{\tau} \tag{1.235}$$

1.3.2.6 Silicon JFET Technologies

The IC fabrication technology used to make JFETs depends primarily on the material. Discrete Si JFETs are available that provide f_T above 500 MHz and very low input rms noise currents through optimizing the channel design and minimizing parasitic capacitances, resistances, and gate diode leakage currents.

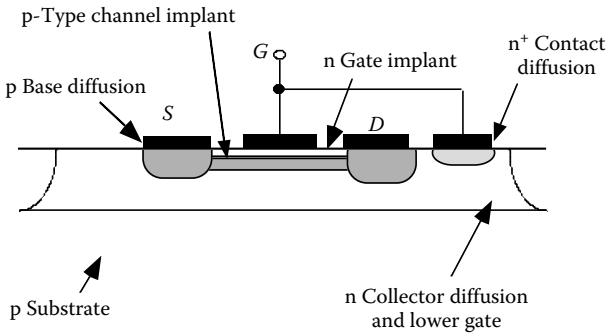


FIGURE 1.54 Cross section of an ion implanted silicon JFET (not to scale).

However, a silicon IC process is rarely designed to optimize the performance of the JFET; rather, the JFET is made to accommodate an existing bipolar process with as few modifications as possible [10]. Then, the extra circuit design flexibility and performance benefits of a relatively inexpensive mixed FET/BJT process (often called BiFET) can be obtained with small incremental cost.

In principle, it would be possible to build p-channel Si JFETs in a standard analog BJT process without additional mask steps if the base diffusion had suitable doping and thickness to give a useful pinch-off voltage when overlaid with the emitter diffusion. Unfortunately, this is usually not the case, since the emitter diffusion is too shallow, and the pinch-off voltage resulting from this approach would be too high (positive in the case of the p-channel device). Therefore, the channel of the JFET must be made thinner either through the use of an additional diffusion or by providing the channel and gate with ion implantations.

In analog ICs applications, silicon JFETs are passengers on a bipolar process; they must be compatible with the BJT process that they inhabit. Most flexibility in the JFET design is achieved using the ion implantation method. Figure 1.54 illustrates the cross section of an ion implanted JFET. In order to gain good control of the pinch-off voltage and transconductance, both the channel and the gate are formed by ion implantation. In addition, the forced compatibility with the BJT process requires use of the collector layer under the channel. This forms a lower gate electrode which is less heavily doped than the channel. Therefore, the depletion region at this interface extends primarily into the collector region, and the lower gate is less effective in contributing to the total transconductance of the JFET. It does add the parasitic capacitance C_{gss} to the device at the collector to substrate junction, limiting frequency response. In addition, the predeposition of channel and gate charge is much more repeatable with ion implantation than with earlier double diffusion methods, so device matching and reproducibility of pinch-off voltage is greatly improved. The f_T will be improved by the larger g_m per unit width and the slightly reduced gate capacitances, and the drain breakdown voltage will be increased as is often needed for an analog IC process. However, low-channel doping is not a good recipe for a high-frequency transistor with short gate length, so the f_T of these devices is still only 50 MHz or so.

1.3.3 Compound Semiconductor FET Technologies

An introduction to compound semiconductor materials will be presented in this section to establish the underlying rationale for using these materials for extremely high-performance MMIC and RFIC applications. The transport properties of typical III–V materials are compared with silicon and SiGe alloys.

There is no denying that silicon is the workhorse of the semiconductor industry. Large, high-quality substrates are relatively inexpensive, a highly stable oxide can be grown with low interface state density, and a highly advanced processing technology has enabled extremely large circuit density and extremely

fine lines to be achieved with low parasitic capacitances. Its greatest weakness for electronic device applications is the relatively low electron velocity and mobility. These intrinsic properties lead to higher transit times and access resistances, respectively, a limitation on high frequency device performance. The deeply scaled submicron technology has compensated for these deficiencies to some degree by aggressive reduction in gate length or base width. Also, p-SiGe has higher hole mobility than p-Si, so access resistance can be improved. And, using the strain induced by local depositions of SiGe in MOSFETs increases electron and hole mobilities. As good as Si IC technology is, there exist compound semiconductor materials whose intrinsic electron velocity and mobility are greatly superior to Si and so can potentially offer higher frequency, higher speed or higher power performance.

The III–V FET and bipolar device technology can provide the highest frequency and lowest noise circuit applications. Its main limitation is density. Device footprints are often significantly larger than those of similar Si devices. Thus, the high intrinsic performance of these devices is achieved in circuits of relatively low complexity.

1.3.3.1 Defining III–V Compound Semiconductors

The compound semiconductor family, as traditionally defined, is composed of the group III and group V elements shown in Table 1.3 [12]. Each semiconductor is formed from at least one group III and one group V element.

The main motivation for using the III–V compound semiconductors for device applications is found in their electronic properties when compared with those of the dominant semiconductor material, silicon. Figure 1.55 is a plot of steady-state electron velocity of several n-type semiconductors versus electric

TABLE 1.3 The Group II–VI Elements

II	III	IV	V	VI
Be	B	C	N	O
Mg	Al	Si	P	S
Zn	Ga	Ge	As	Se
Cd	In	Sn	Sb	Te

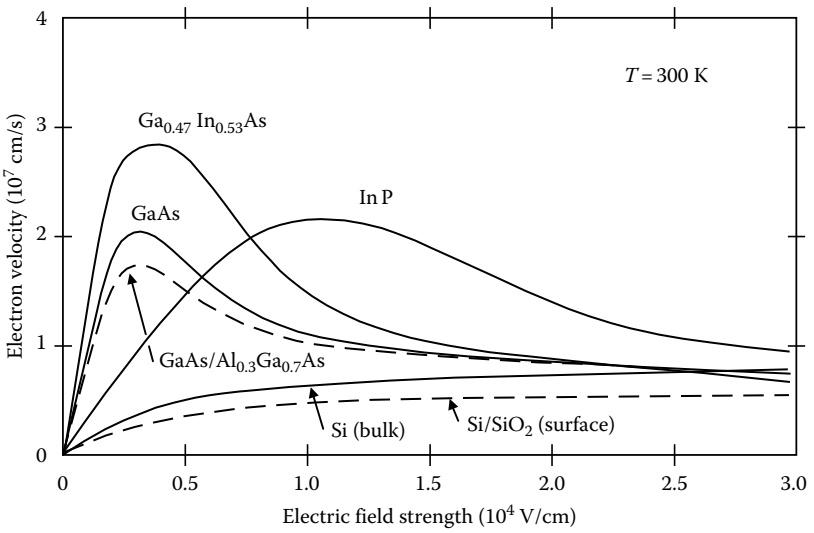


FIGURE 1.55 Electron velocity versus electric field for several n-type semiconductors.

field [12]. From this graph, we see that at low electric fields the slope of the III–V semiconductor curves (mobility) is higher than that of silicon. High mobility means that the semiconductor resistivity will be less for III–V n-type materials, and it therefore will be easier to achieve lower access resistance. Access resistance is the series resistance between the device contacts and the internal active region. An example would be the base resistance of a bipolar transistor or source resistance of a FET. Lower resistance will reduce some of the fundamental device time constants that often dominate device high frequency performance. Figure 1.55 also shows that the peak electron velocity is higher for the III–V's, and the peak velocity can be achieved at much lower electric fields. High velocity reduces transit time, the time required for a charge carrier to travel from its source to its destination, and improves device high-frequency performance. Achieving this high velocity at lower electric fields means that the devices will reach their peak performance at lower voltages, useful for low power, high-speed applications. Higher velocity of electrons also increases the current density of a device since current is the product of charge and velocity. Mobility and peak velocities of several semiconductors are compared in Table 1.4 [12].

The higher velocities are a consequence of the band structure of III–V materials. Since Si is an indirect bandgap material, conduction electrons reside in a high effective mass conduction band (CB). Mobility is dominated by the high effective mass. At high electric fields, the optical phonon generation process limits the maximum achievable electron drift velocity. GaAs, on the other hand, is direct gap, the electron mobility is high because of the lower energy, low effective mass CB where conduction electrons are confined at low fields. However, the average electron velocity will be reduced at higher electric fields due to scattering into the higher mass CB. This produces a saturated drift velocity less than the peak drift velocity, typical of the direct-gap III–V's.

To obtain significant transit velocity improvement over silicon, one must use a ternary III–III–V semiconductor such as InGaAs. The high effective mass CB is separated by 50% of the bandgap for InGaAs, whereas for GaAs it was only 20%. Thus, the peak velocity in InGaAs can be much higher than GaAs because more energy can be transferred to the conduction electrons before they begin scattering to the high mass CB. This results in higher peak velocity, 2.7×10^7 cm/s vs. 2×10^7 cm/s for GaAs.

Also shown in Table 1.4, p-type III–V semiconductors have rather poor hole mobility when compared with elemental semiconductor materials such as silicon or germanium. Holes also reach their peak velocities at much higher electric fields than electrons. Consequently, there has been very little use of p-channel III–V FET devices. The only reason to use compound semiconductor FETs is their superb high-frequency performance. The p-channel devices cannot provide this.

TABLE 1.4 Electronic Properties of Compound Semiconductors Compared with Si and Ge

Semiconductor	E_G (eV)	ε_r	Electron Mobility (cm^2/V-s)	Hole Mobility (cm^2/V-s)	Peak Electron Velocity (cm/s)
Si (bulk)	1.12	11.7	1,450	450	NA
Ge	0.66	15.8	3,900	1,900	NA
InP	1.35 D	12.4	4,600	150	2.1×10^7
GaAs	1.42 D	13.1	8,500	400	2×10^7
$Ga_{0.47}In_{0.53}As$	0.78 D	13.9	11,000	200	2.7×10^7
InAs	0.35 D	14.6	22,600	460	4×10^7
$Al_{0.3}Ga_{0.7}As$	1.80 D	12.2	1,000	100	—
AlAs	2.17	10.1	280	—	—
$Al_{0.48}In_{0.52}As$	1.92 D	12.3	800	100	—
GaN	3.39D	9.0	1,500	30	$2.5–2.7 \times 10^7$
SiC (4H)	3.26	9.8	500		2.2×10^7

Note: In bandgap energy column the symbol "D" indicates direct bandgap, otherwise it is indirect bandgap.

1.3.3.2 Heterojunctions

Heterojunctions provide an additional degree of freedom that is widely used to improve performance of compound semiconductor FET devices. The heterojunction formed by an atomically abrupt transition between AlGaAs and GaAs, shown in the energy band diagram of Figure 1.56 [12], creates discontinuities in the valence and CBs. The CB energy discontinuity is labeled ΔE_C and the valence band discontinuity, ΔE_V. Their sum equals the energy bandgap difference between the two materials. The potential energy steps caused by these discontinuities are used as barriers to electrons or holes. The relative sizes of these potential barriers depend on the composition of the semiconductor materials on each side of the heterojunction. In this example, an electron barrier in the CB is used to confine carriers into a narrow potential energy well with triangular shape. Quantum well structures such as these are used to improve device performance through two-dimensional charge transport channels, similar to the role played by the inversion layer in MOS devices. The structure and operation of heterojunctions in FETs will be described in Section 1.3.3.

The overall principle of the use of heterojunctions is summarized in a *Central Design Principle*:

> Heterostructures use energy gap variations in addition to electric fields as forces acting on holes and electrons to control their distribution and flow [13,14].

The energy barriers can control motion of charge both across the heterojunction and in the plane of the heterojunction. In addition, heterojunctions are most widely used in light emitting devices since the compositional differences also lead to either stepped or graded index of refraction, which can be used to confine, refract, and reflect light. The barriers also control the transport of holes and electrons in the light generating regions.

Figure 1.57 shows a plot of bandgap versus lattice constant for many of the III–V semiconductors [12]. Consider GaAs as an example. GaAs and AlAs have the same lattice constant (approximately 0.56 nm) but different band gaps (1.4 and 2.2 eV, respectively). An alloy semiconductor, AlGaAs, can be grown epitaxially on a GaAs substrate wafer using standard growth techniques. The composition can be selected by the Al to Ga ratio giving a bandgap that can be chosen across the entire range from GaAs to AlAs. Since both lattice constants are essentially the same, very low lattice mismatch can be achieved for any composition of $Al_xGa_{1-x}As$. Lattice matching permits low defect density, high quality materials to be grown that have good electronic and optical properties.

It quickly becomes apparent from Figure 1.57, however, that a requirement for lattice matching to the substrate greatly restricts the combinations of materials available to the device designer. For electron devices, the low mismatch GaAs/AlAs alloys, GaSb/AlSb alloys, and ternary combinations $GaAs/Ga_{0.49}In_{0.51}P$ and $InP/In_{0.53}Ga_{0.47}As/In_{0.52}Al_{0.48}As$ alone are available. Efforts to utilize

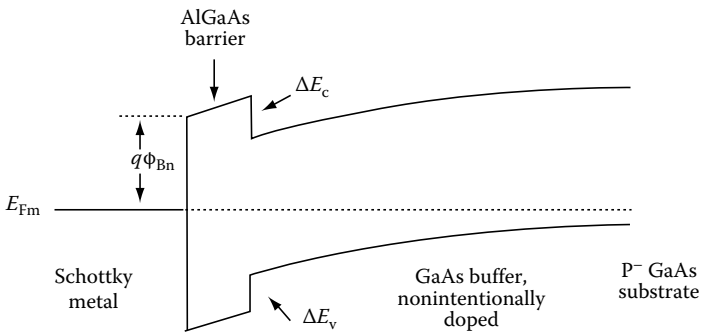


FIGURE 1.56 Energy band diagram of an abrupt heterojunction. Typical AlGaAs/GaAs HEMT band diagram.

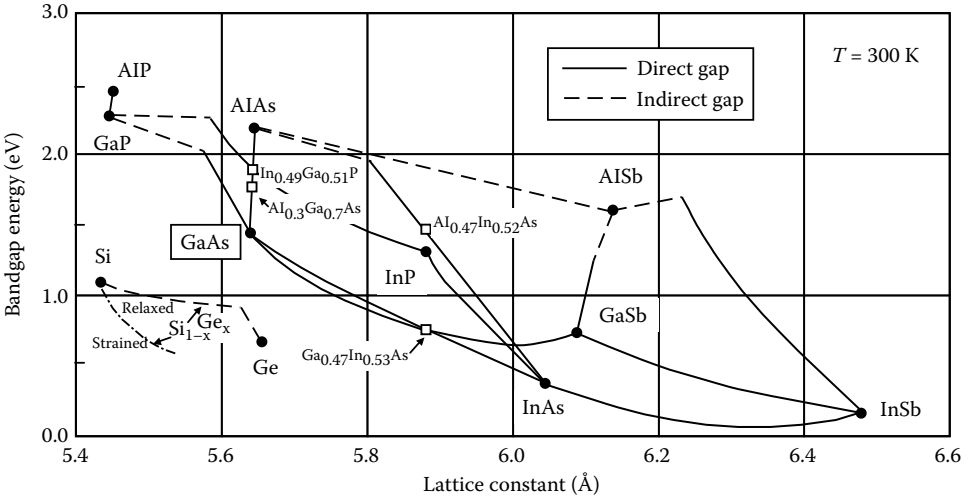


FIGURE 1.57 Energy bandgap versus lattice constant for compound semiconductor materials.

combinations such as GaP on Si or GaAs on Ge that lattice match have been generally unsuccessful because of problems with interface structure, polarization, and autodoping.

For several years, lattice matching was considered to be a necessary condition if mobility-damaging defects were to be avoided. This barrier was later broken when it was discovered that high quality semiconductor materials could still be obtained although lattice-mismatched if the thickness of the mismatched layer is sufficiently small [15,16]. This technique, called pseudomorphic growth, opened another dimension in III–V device technology, and allowed device structures to be optimized over a wider range of bandgap for better electron or hole dynamics and optical properties.

Two of the pseudomorphic systems that have been very successful in high performance millimeter-wave FETs are the InAlAs/InGaAs/GaAs and InAlAs/InGaAs/InP systems. The $In_xGa_{1-x}As$ layer is responsible for the high electron mobility and velocity which both improve as the In concentration x is increased. Up to $x = 0.25$ for GaAs substrates and $x = 0.80$ for InP substrates have been demonstrated and result in great performance enhancements when compared with lattice-matched combinations. [6]

InP substrates, however, are more expensive, smaller, and more easily broken than GaAs. And, the 3.8% lattice mismatch would seem to be too great for direct epitaxy of $In_{0.53}Ga_{0.47}As$ on GaAs substrates. It has been demonstrated, however, that good quality devices can be obtained using the metamorphic growth technique. A thick InP transition layer or a graded InGaP layer is grown directly upon a GaAs substrate. The defects caused by the lattice mismatch are largely contained in this layer, and low defect layers can be obtained when grown upon this transitional buffer layer [17,18].

1.3.3.3 Compound Semiconductor HEMT Devices

High performance GaAs MESFET* and HEMT† devices are constructed with a metal-to-semiconductor junction gate instead of a diffused or implanted pn junction gate. The metal gate forms a Schottky barrier diode directly on an n-type channel or on a wider bandgap barrier layer. In the case of the MESFET as shown in Figure 1.58 [19], the gate, directly on the n-type doped channel, forms a depletion layer which allows the channel height to be varied in the same manner as the JFET. No gate dielectric or p-type diffusion is necessary. Often the gate is deposited in a recess, etched below the surface of the channel. This allows for thicker and sometimes more highly doped regions at source and drain to be used to

* Metal–semiconductor FET.
† High electron mobility transistor.

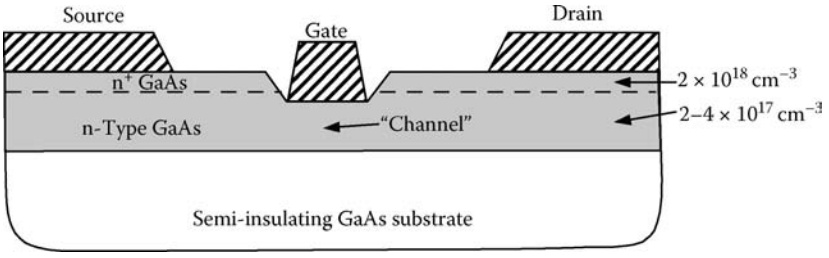


FIGURE 1.58 Cross section of recessed gate GaAs MESFET. (From Estreich, D., in *The VLSI Handbook*, CRC Press, Boca Raton, FL, 2006.)

reduce parasitic resistances. When gate cross sections are very small, for example less than 150 nm, the gate metal is taller than the width, and a thicker, wider region is often deposited on the top to reduce gate access resistance.

With the HEMT device, the gate potential modulates the height of a triangular potential well (Figure 1.56) thereby varying the channel charge available for source–drain conduction. The channel layer is confined by the triangular potential well formed at the interface between the higher bandgap barrier (InAlAs or AlGaAs) and channel (InGaAs or GaAs) as illustrated in the device cross sectional drawing in Figure 1.59 [19]. In some devices, the back side of the channel is also confined by a wide gap barrier. The confinement provided by these energy barriers provides large channel electron sheet concentrations, improving g_m and current density. The active region of the HEMT is formed by epitaxial growth of the channel and barrier region with molecular beam epitaxy.

In Figure 1.59, the device is also shown with a recessed gate. This type of structure enables the use of more highly doped, lower bandgap material on the surface to reduce parasitic source and drain resistances.

These compound semiconductor FETs are used as the primary active device in analog microwave and mm-wave monolithic integrated circuits (MMICs or RFICs). Extremely low noise figure and wide bandwidth have been obtained by the use of HEMT, p-HEMT (pseudomorphic HEMT), and m-HEMT (metamorphic HEMT) devices. These devices achieve their improved performance mainly through the high mobility, undoped InGaAs channel material. The electron velocity vs. electric field of $In_{0.53}Ga_{0.47}As$ is compared with GaAs and Si in Figure 1.55 where it can be seen that higher drift velocity is obtained in $In_{0.53}Ga_{0.47}As$ [20] than either GaAs [21] or Si [22]. The higher the In concentration in the InGaAs, the higher the mobility and velocity and the lower the noise.

Finally, the gate barrier heterojunction also enables good Schottky gate characteristics to be obtained even though the channel material itself has a low bandgap and would otherwise provide a poor barrier height if the metal were in direct contact.

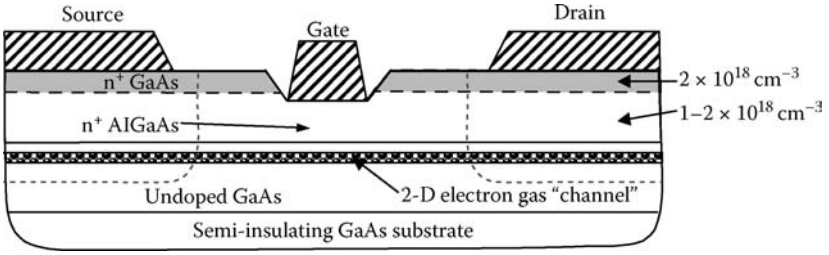


FIGURE 1.59 Cross section of recessed gate AlGaAs/GaAs HEMT device structure. (From Estreich, D., in *The VLSI Handbook*, CRC Press, Boca Raton, FL, 2006.)

TABLE 1.5 Microwave and mm-Wave Performance Comparison between Compound Semiconductor FETs

Frequency (GHz)	Device	Gain (dB)	Noise Figure (dB)	Reference
4–9	Dual Gate	21	<1.75	[1]
10–20	100 nm	17	<2.75	
20–40	GaAs pHEMT	20	<2.5	
23	130 nm	43	1.9	[2]
18–40	InP HEMT	>40	—	
0.5–80	100 nm	>17	<2.5	[3]
	InP HEMT			
70–105	50 nm	20	2.4	[4]
220	m-HEMT	21	9	
192–235	50 nm	>15	—	[5]
	m-HEMT			
270	35 nm	11.6	—	[6]
300	InP p-HEMT	6		

Excellent performance of HEMT, p-HEMT, and m-HEMT MMICs at microwave and millimeter wave frequencies has been reported. Table 1.5 presents a summary of some representative MMIC amplifiers where both narrowband and wideband amplifiers are reported. Gate lengths down to 35 nm have been successfully used for mm-wave and sub-mm-wave amplifiers.

GaAs HEMT devices also exhibit higher breakdown voltages (often 15 to 20 V) that make them suitable for power amplifier applications. In a recent article, a wideband distributed GaAs HEMT amplifier was reported that provided over 4 W of output power from 4 to 18 GHz at a 5 V drain bias with a power added efficiency of 23% [23]. Narrowband HEMT amplifiers can provide much higher efficiency and power.

1.3.3.4 Wide Bandgap Compound Semiconductors

In recent years there has been increasing interest in the wide bandgap compound semiconductors, SiC and GaN (and associated alloys of Al/In/Ga with N). The applications have been primarily for microwave power applications because the wider bandgap increases breakdown voltage while the band structure allows for high electron peak velocities in both materials [24,25]. Table 1.6 compares the fundamental physical properties of the wide-gap compound semiconductors with GaAs and Si [12]. It should be noted that there is not uniform agreement on the wide-gap parameter values from one reference to the next, but the numbers presented are representative of the current literature. As seen in Table 1.6, thermal conductivity is very high for both SiC and GaN, allowing for effective removal of heat from power devices. In fact, at room temperature, SiC has a higher thermal conductivity than any metal.

Figure 1.60 compares the electron velocity of GaN and SiC with GaAs and silicon [12,26]. The peak velocity of GaN is reached at electric fields above 150 kV/cm. Both SiC and GaN retain their good

TABLE 1.6 High Electron Mobility GaN

Material	Bandgap (eV)	Mobility (cm^2/V-s)	E_c (V/cm)	Saturation Drift Velocity (cm/s)	Thermal Conductivity (W/cm-K)
n-SiC (4H)	3.26	500	2.2×10^6	2×10^7	3.0–3.8
n-GaN	3.39	1500	3×10^6	1.5×10^7	2.2
n-GaAs	1.4	5000	3×10^5	0.6×10^7	0.45
n-Si	1.1	1300	2.5×10^5	1×10^7	1.45

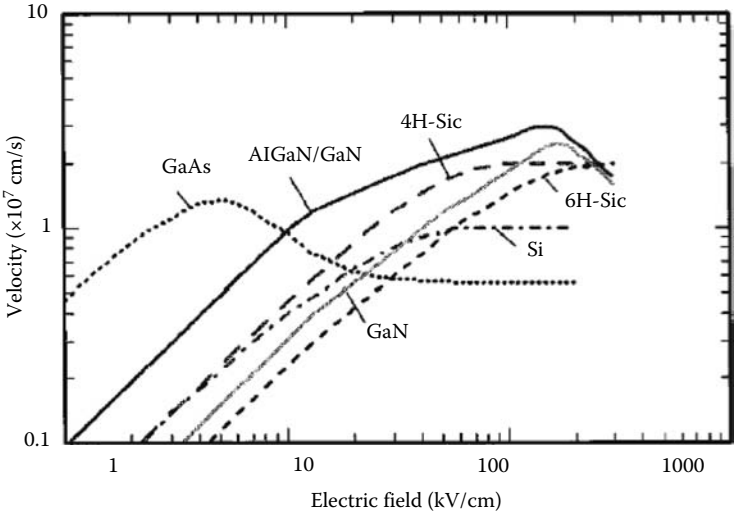


FIGURE 1.60 Electron velocity versus electric field of GaN and SiC compared with Silicon and GaAs. (From Trew, R.J., *Proc. IEEE*, 90, 1032, 2002. With permission.)

transport properties for high power applications. Table 1.6 shows that GaN has high electron mobility as well, which helps to reduce parasitic source resistance. Hole mobility for wide bandgap compound semiconductors is quite low, however, generally less than 50 $cm^2/V-s$.

There is significant lattice mismatch between a GaN channel in a heterojunction FET and the AlGaN barrier layer. However, the strain caused by this mismatch produces polarization and piezoelectric effects that induce large sheet charge densities, above $10^{13}\,cm^{-2}$ in the channel, beneficial for high current density operation of these devices [26]. This level of charge is about five times higher than what can be induced in GaAs channels in the AlGaAs/GaAs heterostructure.

Figures of merit are often employed when comparing materials for microwave power amplifier applications. Johnson's FOM [27]

$$\text{JFOM} = \frac{E_c v_{sat}}{2\pi} \tag{1.236}$$

has units of power–frequency. E_c is the maximum or critical electric field for breakdown and v_{sat} is the saturated drift velocity at high electric fields. This expresses the electronic merits of the material but neglects to consider thermal conductivity, also of importance for power electronics. Nevertheless, based on the electronic properties alone, GaN and SiC have a JFOM approximately 18 times greater than Si or GaAs. If the superior thermal conductivity is also considered, it becomes clear that these materials are extremely well suited for microwave power.

1.3.3.5 GaN HEMT Field Effect Transistors

The unusual electronic and thermal properties of the wide bandgap materials such as GaN are very attractive for applications requiring transistors with both high breakdown voltage and high frequency performance. This range of applications is focused especially on microwave power transistors. To understand why, consider first what is most desirable in a high power transistor's drain current–voltage characteristics. Figure 1.61 shows an idealized representation of this characteristic. To obtain the maximum output sinusoidal voltage and current amplitudes, one would like a device with high breakdown voltage, V_{break}, low "knee" voltage, V_{knee}, and high maximum current, I_{max}. The voltage swing is

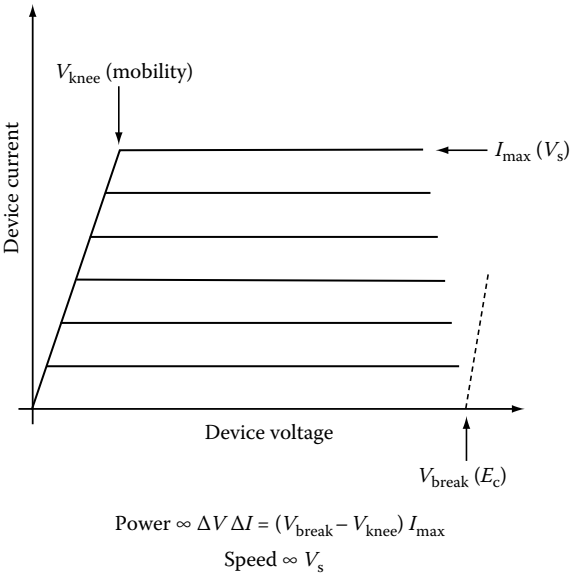


FIGURE 1.61 Idealized GaN HEMT drain current–drain voltage characteristic.

determined by $V_{break} - V_{knee}$, and the current swing by I_{max}. This combination provides high power density (W/mm) in a device, therefore requiring smaller device area. Power is proportional to

$$P_{OUT} \propto (V_{break} - V_{knee}) \times I_{max}.$$

DC to RF conversion efficiency, equally important in a power device, is proportional to

$$\text{Efficiency} \propto (1 - V_{knee}/V_{break}).$$

Thus, a large breakdown voltage is helpful for both power and efficiency.

But, Si LDMOS devices also have high breakdown voltage. So, why is GaN better? First, the current density is far higher in GaN. The high current density in the GaN HEMT is a result of the exceptionally high sheet charge, n_S, in the channel, typically 1×10^{13} cm^{-2} or higher. The high sheet charge density is a result of the static polarization that occurs at the interface between the GaN channel and the AlGaN barrier. To satisfy the charge balance at the interface, a high density of negative charge is required. A cross section of a GaN HEMT device is shown in Figure 1.62.

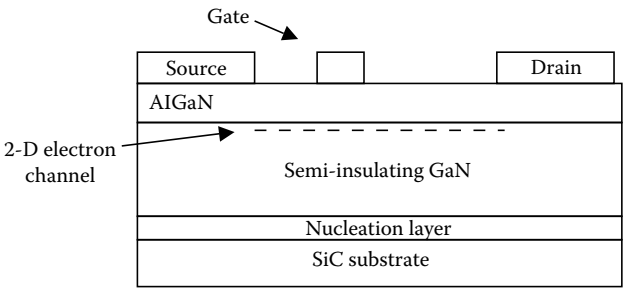


FIGURE 1.62 Cross section of GaN HEMT device structure.

Second, the high current and power density means that smaller device areas can be used to meet a particular power requirement. Smaller area translates into higher impedances because capacitances are proportional to device area. This simplifies the matching at the input and output, and therefore can lead to wider bandwidth and lower losses. The GaN devices are grown on a semi-insulating SiC substrate, thus the drain–source capacitance, C_{ds}, is small: typically about 0.25 pF/mm of device width. A 25 W device would require about 5 mm of channel width giving C_{ds} of about 1.25 pF. The Si device drain capacitance is generally quite high by comparison. For example, a 25 W LDMOS device designed for operation at 500 MHz has a C_{ds} of about 30 pF. This greatly limits the application of such devices for higher frequency or switching mode PA applications.

The current density and high frequency performance are also aided by the high saturated electron velocity in GaN. Refer once again to Figure 1.60 where electron velocity at high electric field in GaN is compared with other semiconductor materials. The electron transit time across the channel is inversely proportional to the carrier velocity; the current directly proportional.

The GaN HEMT gate structure often includes a field plate at the drain end of the gate as shown in Figure 1.63. The field plate distributes the electric field in the gate–drain region over a wider area leading to reduction in peak electric field and higher breakdown voltage. In addition, surface trap charge is less affected by potential variation on the surface, thus there is less depletion of channel charge by these traps leading to reduction in the low frequency dispersion effects.

Recent GaN HEMT microwave power amplifier performance highlights are presented in Table 1.7. Power outputs of several hundreds of watts have been obtained under low duty cycle pulsed operating conditions where thermal effects are less serious. CW output powers in the same range have been obtained by power combining of two or more amplifiers. Operation at 35 GHz was reported on MMIC PAs using devices with reduced gate length (0.15 μm) and smaller drain voltages.

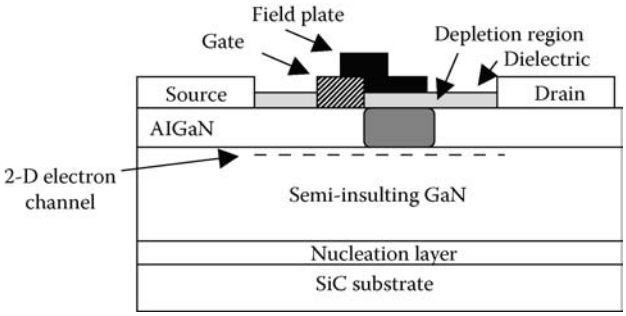


FIGURE 1.63 Cross section of GaN HEMT device structure with gate connected field plate.

TABLE 1.7 Recent GaN/HEMT Microwave Power Amplifier Performance Highlights

F (GHz)	P_{OUT} (W)	PAE (%)	V_{DC} (V)	W/mm	Reference
1.5	500[a]	49	65	13.9[a]	[28]
2.14	750[a]	—	50	7.8[a]	[29]
2.14	370	—	45	3.8	[30]
6	130[a]	45	50[a]	5.4[a]	[31]
9.5	80	34	30	3.5	[32]
35	4[b]	23	24	3.3	[33]
35	0.9[b]	51	20	4.5	[34]

[a] Pulsed CW unless otherwise noted.
[b] MMIC.

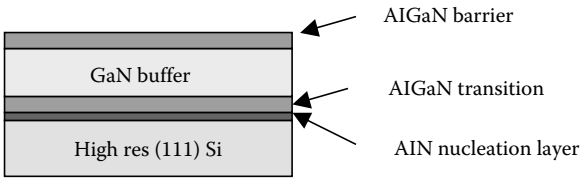


FIGURE 1.64 GaN on Si m-HEMT substrate.

GaN has also been successfully grown on high resistivity silicon substrates as illustrated by Figure 1.64. A thick nucleation layer leads to a metamorphic structure. Good power amplifier GaN on Si HEMT devices have been demonstrated to provide comparable performance in power and efficiency to those grown on Si [35,36]. The thermal conductivity of silicon is considerably less than that of SiC, however, so one would not expect the thermal resistance of the GaN on Si HEMT to be as low as the former.

1.3.4 Conclusion

While the mainstream semiconductor device and circuit technology is defined by silicon and its related materials, the superior electron transport properties of compound semiconductor materials offer unique advantages in applications requiring the highest frequency, speed and power or lowest noise figure. The range of device structural possibilities with compound semiconductor heterojunctions is far greater than what can be realized without this option, and this has led to high-performance FET structures, HEMT, p-HEMT, and m-HEMT, with excellent bandwidth, noise figure, and power.

References

1. Deal, W. R. et al., Design and analysis of broadband dual-gate balanced low-noise amplifiers, *IEEE J. Solid State Circuits*, 42(10), 2107–2115, Oct. 2007.
2. Matsuda, S. et al., Very compact high-gain broadband low-noise amplifier in InP HEMT technology, *IEEE Trans. Microwave Theory Tech*., 54(12) part 2, 4565–4571, Dec. 2006.
3. Grundbacher, R., et al., High performance millimeter wave 0.1/spl mu/m InP HEMT MMIC LNAs fabricated on 100 mm wafers, *International Conference on Indium Phosphide and Related Materials, 16th IPRM*, May 31–June 4; Kagoshima, Japan, pp. 284–287, 2004.
4. Schlechtweg, M. and A. Tessmann, From 100 GHz to terahertz electronics—activities in Europe, *IEEE Compound Semiconductor IC Symposium*, pp. 8–11, San Antonio, TX, Nov. 2006.
5. Leuther, A., et al., 50 nm MHEMT technology for G- and H-Band MMICs, *International Conference on Indium Phosphide and Related Materials, 19th IPRM*, 14–18 May, Matsue Japan, pp. 24–27, 2007.
6. Deal, W. R., et al., Demonstration of a 270-GHz MMIC amplifier using 35-nm InP HEMT technology, *IEEE Microwave Wireless Components Lett*., 17(5), 391–393, May 2007.
7. Sze, S. M., *Physics of Semiconductor Devices*. Wiley-Interscience, New York, 1981.
8. Taylor, G. W., H. M. Darley, et al., A device model for an ion-implanted MESFET. *IEEE Trans. Electron. Dev*., ED-26, 172–179, 1979.
9. Lee, S. J. and C. P. Lee, Temperature effect on low threshold voltage ion-implanted GaAs MESFETs. *Electron. Lett*., 17(20), 760–761, 1981.
10. Gray, P. R. and R. Meyer, *Analysis and Design of Analog Integrated Circuits*. 4th ed., John Wiley, New York, 2004.
11. Statz, H., P. Newman, et al., GaAs FET device and circuit simulation in SPICE, *IEEE Trans. Electron. Dev*., 34(2), 160–169, 1987.
12. Long, S. I., Compound semiconductor materials, Chap. 71 in *The VLSI Handbook*, 2nd edn., Ed. W.-K Chen, CRC Press, Boca Raton, FL, 2006.

13. Kroemer, H., Heterostructures for everything: Device principles of the 1980's? *Japanese J. Appl. Phys*., 20, 9, 1981.
14. Kroemer, H., Heterostructure bipolar transistors and integrated circuits, *Proc. IEEE*, 70, 13, 1982.
15. Matthews, J. W. and A. E. Blakeslee, Defects in epitaxial multilayers, III. Preparation of almost perfect layers, *J. Crystal Growth*, 32, 265, 1976.
16. Matthews, J. W. and A. E. Blakeslee, Coherent strain in epitaxially grown films, *J. Crystal Growth*, 27, 118, 1974.
17. Hoke, W. E., et al., Properties of metamorphic materials and device molecular beam epitaxy, *2002 International Conference on Molecular Beam Epitaxy*, 15–20 September, San Francisco, CA, pp. 69–70, 2002.
18. Schlechtweg, M., et al., Millimeter-wave and mixed-signal integrated circuits based on advanced metamorphic HEMT technology, *16th International Conference on Indium Phosphide and Related Materials*, May 31–June 4; Kagoshima, Japan, pp. 609–614, 2004.
19. Estreich, D., Compound semiconductor devices for analog and digital circuits, Chap. 72 in *The VLSI Handbook*, 2nd edn., Ed. W.-K Chen, CRC Press, Boca Raton, FL, 2006.
20. Littlejohn, M. A., et al., High-field transport in InGaAs and related heterostructures, in *Properties of Lattice-Matched and Strained Indium Gallium Arsenide*, Inspec—IEE, London, pp. 107–116, 1993.
21. Pozela, F. and A. Reklaitis, Electron transport properties in GaAs at high electric fields, *Solid-State Elect*., 23, 927–933, 1980.
22. Jacobini, C., et al., A review of some charge transport properties of Silicon, *Solid-State Elect*., 20, 77, 1977.
23. Meharry, D., et al., Multi-watt wideband MMICs in GaN and GaAs, *IEEE/MTT-S International Microwave Symposium*, June 3–June 8; Honolulu; Hawaii, pp. 631–634, 2007.
24. Gelmont, B., K. Kim, and M. Shur, Monte-Carlo simulation of electron transport in gallium nitride, *J. Appl. Phys*., 74, 1818–1821, 1993.
25. F. Schwierz, An electron mobility model for wurtzite GaN, *Solid State Elect*., 49, 889–895, 2005.
26. Trew, R. J., SiC and GaN transistors: Is there one winner for microwave power applications? *Proc. IEEE*, *Special Issue on Wide Bandgap Semiconductors*, 90, 1032–1047, June 2002
27. Johnson, E. O., Physical limitations on frequency and power parameters of transistors, *IRE International Convention Record*, Vol. 13; Part 5, pp. 27–34, March 1965.
28. Maekawa, A., et al., A 500W Push-Pull AlGaN/GaN HEMT amplifier for L band high power applications, *IEEE/MTT-S International Microwave Symposium*, IEEE Mircrowave Theory and Techniques Society, San Francisco, CA, June 11–16, pp. 722–725, 2006.
29. Wakejima, A., et al., Pulsed 0.75 kW output single-ended GaN-FET amplifier for L/S band applications, *Electron. Lett*., 42, 1349–1350, Nov. 2006.
30. Wakejima, A., et al., 370 W output power GaN-FET amplifier for W-CDMA cellular base stations, *Electron. Lett*., 41, 1371–1372, Dec. 2005.
31. Yamanaka, K., et al., C-band GaN HEMT power amplifier with 220W output power, *2007 IEEE International Microwave Symposium*, June 8, Honolulu, Hawaii, pp. 1251–1254, 2007.
32. Takagi, K., et al., Xband AlGaN/GaN HEMT with over 80 W output power, *IEEE Compound Semiconductor IC Symposium* (*CSICS*), Nov 12–15, San Antonio, TX, pp. 265–268, 2006.
33. Darwish, A., et al., 4 W Ka-band AlGaN/GaN power amplifier MMIC, *IEEE/MTT-S International Microwave Symposium*, IEEE Microwave Theory and Techniques Society, San Francisco, CA, June 11–16, pp. 730–733, 2006.
34. Kao, M.-Y., et al., AlGaN/GaN HEMTs with PAE of 53% at 35 GHz for HPA and multi-function MMIC applications, *2007 IEEE International Microwave Symposium*, Honolulu, Hawaii, June 3–8, pp. 627–629, 2007.
35. Johnson, J. W., Piner, E. L., Vescan, A., Therrien, R., Rajagopal, P., Roberts, J. C., Brown, J. D., Singhal, S., and Linthieum, K. J., 12 W/mm AlGaN/GaN HFETs on silicon substrates, *IEEE Elect. Dev. Lett*., 25(7), 459–461, July 2004.
36. Nagy, W. et al., 150 W GaN on Si RF power transistor, *IEEE/MIT-S International Microwave Symposium*, Microwave Theory and Techniques Society, Long Beach, CA, 4pp, June 2005.

1.4 Passive Components

Nhat M. Nguyen

1.4.1 Resistors

Resistors available in monolithic form are classified in general as semiconductor resistors and thin-film resistors. Semiconductor structures include diffused, pinched, epitaxial, and ion-implanted resistors. Commonly used thin-film resistors include tantalum, nickel–chromium (Ni–Cr), cermet (Cr–SiO), and tin oxide (SnO_2). Diffused, pinched, and epitaxial resistors can be fabricated along with other circuit elements without any additional processing steps. Ion-implanted and thin-film resistors require additional processing steps for monolithic integration but offer lower temperature coefficient, smaller absolute value variation, and superior high-frequency performance.

Resistor calculation. The simplified structure of a uniformly doped resistor of length L, width W, and thickness T is shown in Figure 1.65. The resistance is

$$R = \frac{1}{\sigma}\frac{L}{WT} = \left(\frac{\rho}{T}\right)\frac{L}{W} = R_{\mathrm{n}}\frac{L}{W} \tag{1.237}$$

where

σ and ρ are conductivity and resistivity of the sample, respectively
R_{N} is referred to as the *sheet resistance*

From the theory of semiconductor physics, the conductivity of a semiconductor sample is

$$\sigma = q\left(\mu_n n + \mu_p p\right) \tag{1.238}$$

where

q is the electron charge (1.6×10^{-19} C)
$\mu_n(\mathrm{cm}^2/\mathrm{V}\cdot\mathrm{s})$ is the electron mobility
$\mu_p(\mathrm{cm}^2/\mathrm{V}\cdot\mathrm{s})$ is the hole mobility
$n(\mathrm{cm}^{-3})$ is the electron concentration
$p(\mathrm{cm}^{-3})$ is the hole concentration
$\sigma(\Omega/\mathrm{cm})^{-1}$ is the electrical conductivity

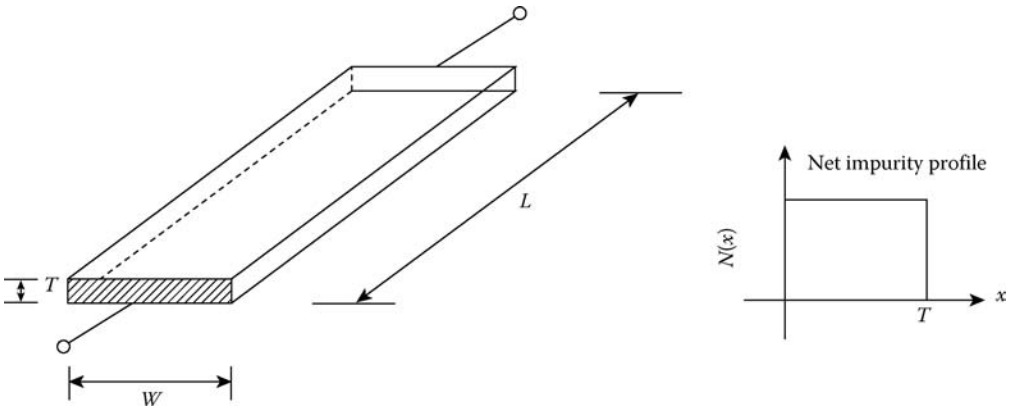


FIGURE 1.65 Simplified structure of a uniformly doped resistor.

For an n-type doped sample with a concentration $N_D(\text{cm}^{-3})$ of donor impurity atoms, the electron concentration n is approximately equal to N_D. Given the mass-action law $np = n_i^2$, the conductivity of an n-type doped sample is approximated by

$$\sigma = q\left(\mu_n N_D + \mu_p \frac{n_i^2}{N_D}\right) \approx q\mu_n N_D \tag{1.239}$$

where $n_i(\text{cm}^{-3})$ is the *intrinsic* concentration. For a p-type doped sample, the conductivity is

$$\sigma = q\left(\mu_n \frac{n_i^2}{N_A} + \mu_p N_A\right) \approx q\mu_p N_A \tag{1.240}$$

where $N_A(\text{cm}^{-3})$ is the concentration of p-type donor impurity atoms. The sheet resistance of an n-type uniformly doped resistor is thus

$$R_n = \left(\frac{1}{q\mu_n N_D T}\right) \tag{1.241}$$

For an n-type nonuniformly doped resistor as shown in Figure 1.66, where n-type impurity atoms are introduced into the p-type region by means of a high-temperature diffusion process, the sheet resistance [7] is

$$R_n = \left[\int_0^{x_j} q\mu_n N_D(x)\mathrm{d}x\right]^{-1} \tag{1.242}$$

where x_j is the distance from the surface to the edge of the junction depletion layer.

Measured values of electron mobility and hole mobility in silicon material as a function of impurity concentration are shown in Figure 1.67 [4]. The resistivity ρ (Ω-cm) of n-type and p-type silicon as a function of impurity concentration is shown in Figure 1.68 [12].

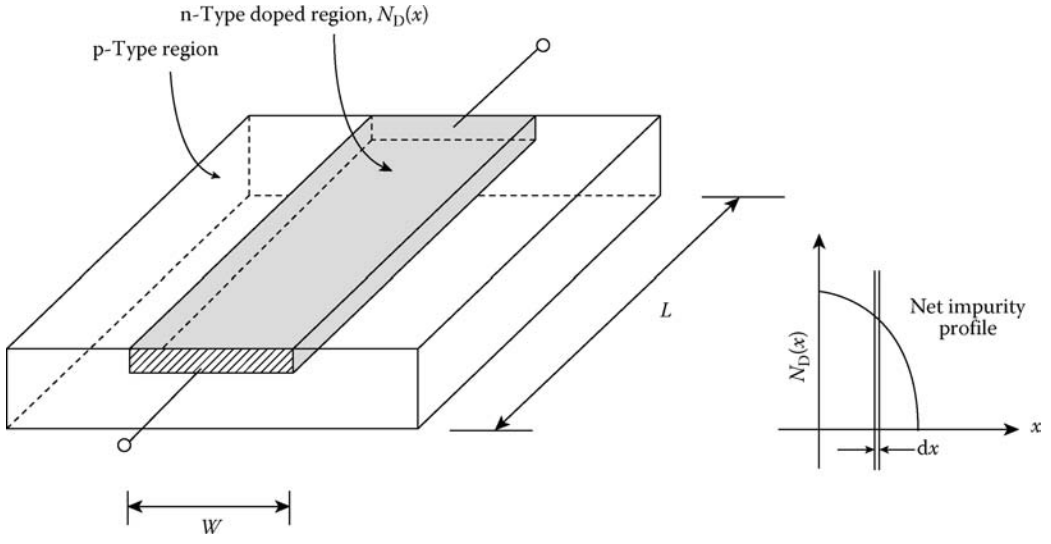


FIGURE 1.66 Simplified structure of an n-type nonuniformly doped resistor.

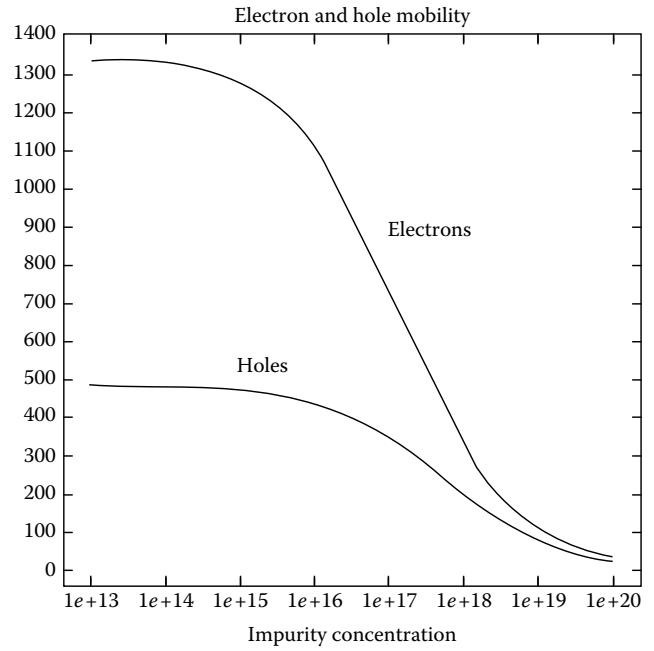


FIGURE 1.67 Electron and hole mobility vs. impurity concentration in silicon.

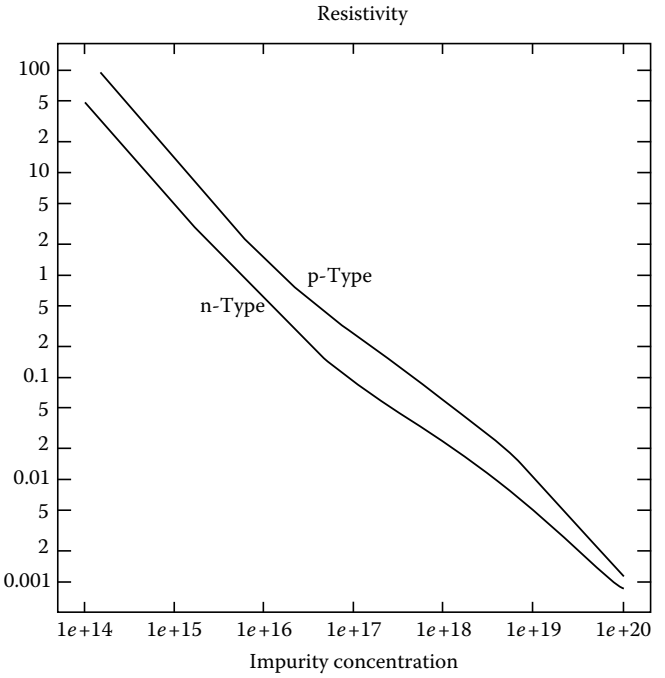


FIGURE 1.68 Resistivity of p-type and n-type silicon vs. impurity concentration.

The sheet resistance depends also on temperature since both electron mobility and hole mobility vary with temperature [17]. This effect is accounted for by utilizing a temperature coefficient quantity that measures the sheet resistance variation as a function of temperature. A mathematical model of the temperature effect is

$$R_v(T) = R_v(T_o)[(T - T_o)\text{TC}] \tag{1.243}$$

where

T_o is the room temperature

"TC" is the temperature coefficient

1.4.1.1 Diffused Resistors

In metal-oxide-semiconductor (MOS) technology, the diffused layer forming the source and drain of the MOS transistors can be used to form a diffused resistor. In silicon bipolar technology, the available diffused layers are base diffusion, emitter diffusion, active base region, and epitaxial layer.

Base-diffused resistors. The structure of a typical base diffused resistor is shown in Figure 1.69, where the substrate material is assumed of p-type silicon material. The diffused resistor is formed by using the p-type base diffusion of the npn transistors. The resistor contacts are formed by etching selected windows of the SiO_2 passivation layer and depositing thin films of conductive metallic material. The isolation region can be formed with either a p-type doped junction or a trench filled with SiO_2 dielectric material. The pn junction formed by the p-type resistor and the n-type epitaxial (epi) layer must be reverse biased

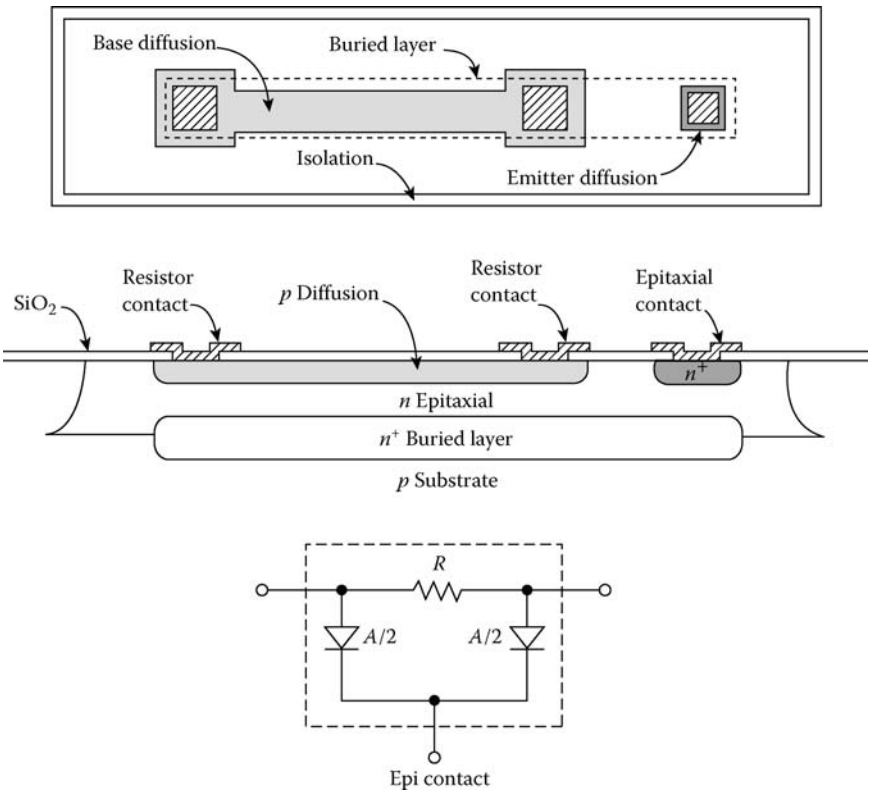


FIGURE 1.69 p-Type base-diffused resistor.

in order to eliminate the undesired dc current path through the pn junction. The impedance associated with a forward-biased pn junction is low and thus would also cause significant ac signal loss. To ensure this reverse bias constraint the epi region must be connected to a potential that is more positive than either end of the resistor contacts. Connecting the epi region to a relatively higher potential also eliminates the conductive action due to the parasitic pnp transistor formed by the p-type resistor, the n-type epi region, and the p-type substrate. When the base-diffused resistor is fabricated along with other circuit elements to form an integrated circuit (IC), the epitaxial contact is normally connected to the most positive supply of the circuit.

The resistance of a diffused resistor is given by Equation 1.237, where the diffused sheet resistance is between 100 and 200 Ω/N. Due to the lateral diffusion of impurity atoms, the effective cross-sectional area of the resistor is larger than the width determined by photomasking. This lateral or side diffusion effect can be accounted for by replacing the resistor width W by an effective width W_{eff}, where $W_{\text{eff}} \geq W$. The resistance from the two resistor contacts must also be accounted for, especially for small values of L/W [3]. Base-diffused resistors have a typical temperature coefficient between +1500 and +2000 ppm/°C.

The maximum allowable voltage for the base-diffused resistor of Figure 1.69 is limited by the breakdown voltage between the p-type base diffusion and the n-type epi. This voltage equals the breakdown voltage BV_{CBO} of the collector–base junction of the npn transistor and typically causes an *avalanche breakdown* mechanism across the base–epi junction. As the applied voltage approaches the breakdown voltage, a large leakage current flows from the epi region to the base region and can cause excessive heat dissipation.

For analog IC applications where good matching tolerance between adjacent resistors is required, the resistor width should be made as large as possible. Base-diffused resistors with 50 μm resistor widths can achieve a matching tolerance of ±0.2%. The minimum resistor width is limited by photolithographic consideration with typical values between 3 and 5 μm. Also, in order to avoid the self-heating problem of the resistor it is important to ensure a minimum resistor width for a given dc current level, with a typical value of about 3 μm for every 1 mA of current.

With respect to high-frequency performance, the reverse-biased pn junction between the p-type base diffusion and the n-type epi contributes a distributed depletion capacitance which in turn causes an impedance roll-off at 20 dB/decade. This capacitance depends on the voltage applied across the junction and the junction impurity-atom dopings. For most applications the electrical lumped model as shown in Figure 1.69 is adequate for characterizing this capacitive effect where the effective pn junction area is divided equally between the two diodes. Figure 1.70 shows a normalized impedance response as a function of the *RC* distributed stage. The frequency at which impedance value is reduced by 3 dB is given by

$$f_{-3\text{ dB}} = \begin{cases} \left(\dfrac{1}{2\pi}\right)\dfrac{2.0}{RC} & N = 1 \text{ (Circuit model of Figure 1.62)} \\ \left(\dfrac{1}{2\pi}\right)\dfrac{2.32}{RC} & N = 2 \\ \left(\dfrac{1}{2\pi}\right)\dfrac{2.42}{RC} & N = 3 \\ \left(\dfrac{1}{2\pi}\right)\dfrac{2.48}{RC} & N = 4 \end{cases} \tag{1.244}$$

Emitter-diffused resistors. Emitter-diffused resistors are formed by using the heavily doped n^+ emitter diffusion layer of the npn transistors. Due to the high doping concentration, the sheet resistance can be as low as 2 to 10 Ω/N with a typical absolute value tolerance of ±20%.

Figure 1.71 shows an emitter-diffused resistor structure where an n^+ diffusion layer is formed directly on top of the n-type epitaxial region and the ohmic contacts are composed of conductive

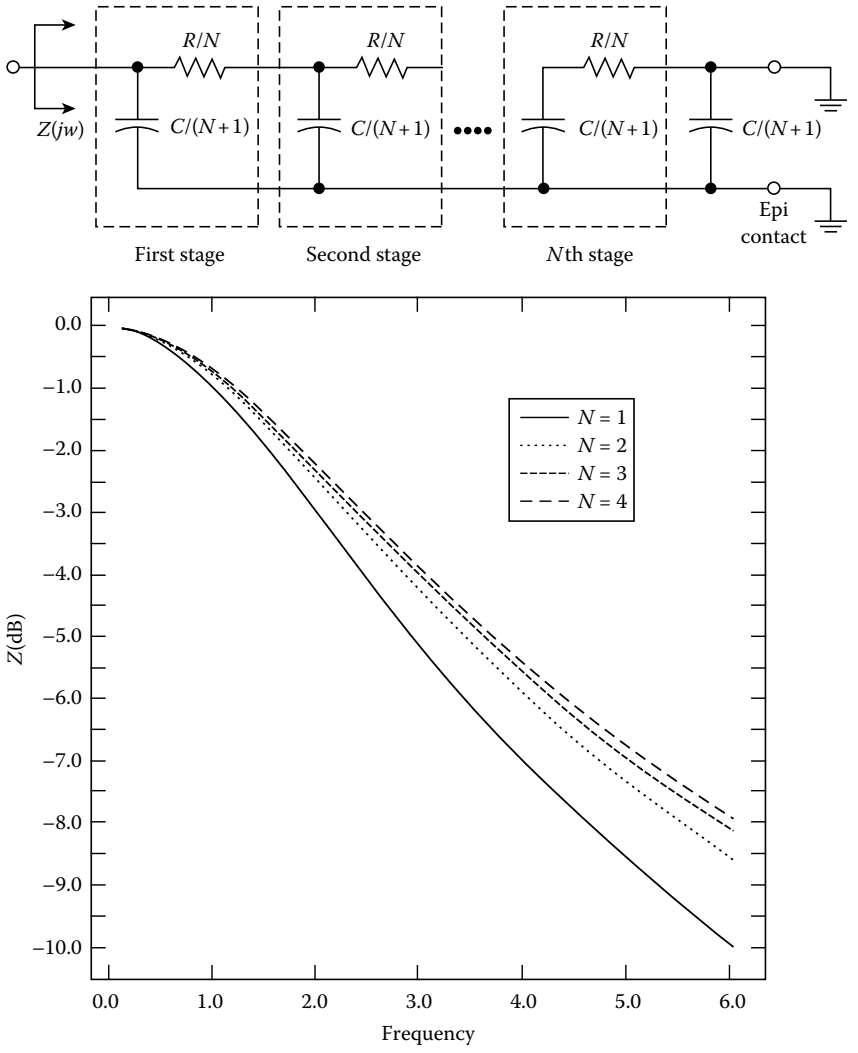


FIGURE 1.70 Normalized frequency response of a diffused resistor for $N = 1, 2, 3, 4$. The epi contact and one end of the resistor are grounded.

metal thin films. Since the resistor body and the epi layer are both n-type doped, they are electrically connected in parallel but the epi layer is of much higher resistivity due to its lower concentration doping, and thus the effective sheet resistance of the resistor structure is determined solely by the n^+ diffusion layer. The pn junction formed between the p-type substrate and the n-type epi region must always be reverse biased, which is accomplished by connecting the substrate to a most negative potential. Because of the common n-type epi layer, each resistor structure of Figure 1.71 requires a separate isolation region.

Figure 1.72 shows another emitter diffused resistor structure where the n^+ diffusion layer is situated within a p-type diffused well. Several such resistors can be fabricated in the same p-type well or in the same isolation region because the resistors are all electrically isolated. The p-type well and the n^+ diffusion region form a pn junction that must always be reverse biased for electrical isolation. In order to eliminate the conductive action due to the parasitic npn transistor formed by the n-type resistor body,

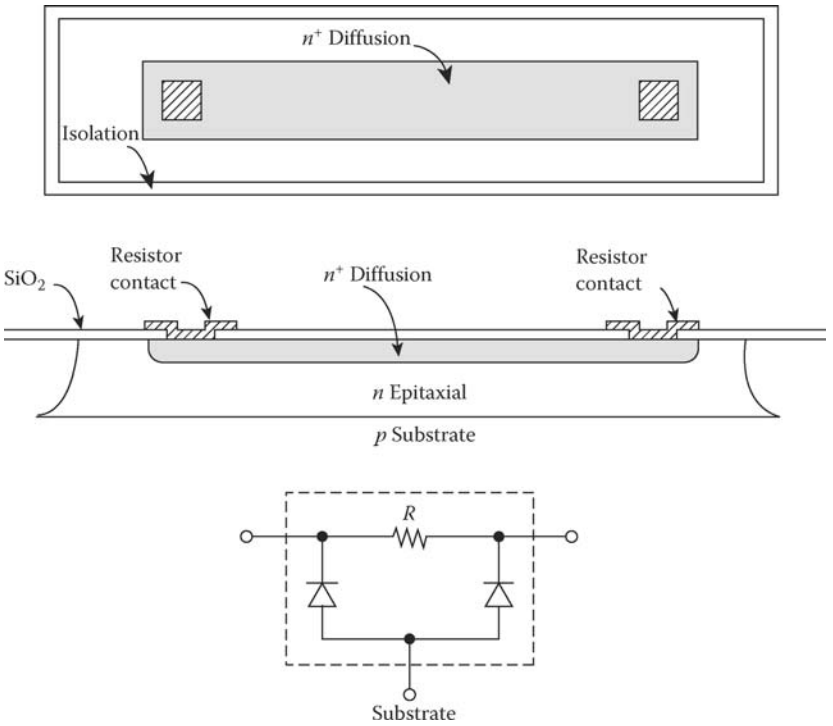


FIGURE 1.71 n-Type emitter-diffused resistor I.

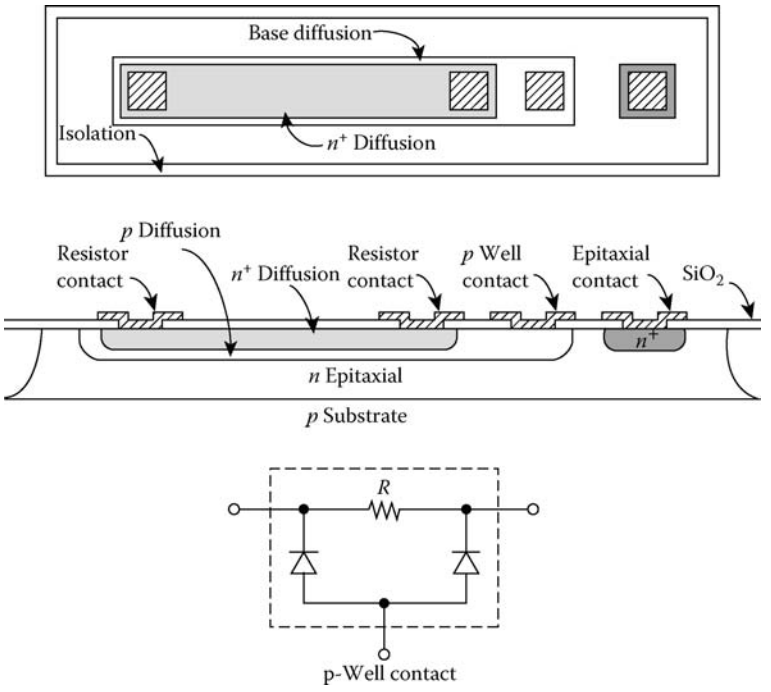


FIGURE 1.72 n-Type emitter-diffused resistor II.

the p-type well, and the n-type epi, the junction potential across the well contact and the epi contact must be either short-circuited or reverse-biased. The maximum voltage that can be applied across the emitter-diffused resistor of Figure 1.72 is limited by the breakdown voltage between the n^+ diffusion and the p-type well. This voltage equals the breakdown voltage BV_{EBO} of the emitter–base junction of the npn transistor, with typical values between 6 and 8 V.

1.4.1.2 Pinched Resistors

The active base region for the npn transistor can be used to construct pinched resistors with typical sheet resistance range from 2 to 10 KΩ/N. These high values can be achieved due to a thin cross-sectional area through which the resistor current traverses. The structure of a p-type base-pinched resistor is shown in Figure 1.73, where the p-type resistor body is ''pinched'' between the n^+ diffusion layer and the n-type epitaxial layer. The n^+ diffusion layer overlaps the p-type diffusion layer and is therefore electrically connected to the n-type epi. In many aspects the base-pinched resistor behaves like a p-channel JFET, in which the active base region functions as the p-channel, the two resistor contacts assume the drain and source, and the n^+ diffusion and the epi constitute the n-type gate. When the pn junction formed between the active base and the surrounding n^+ diffusion and n-epi is subject to a reverse bias potential, the carrier-free depletion region increases and extends into the active base region, effectively reducing the resistor cross section and consequently increasing the sheet resistance. Since the carrier-free depletion region varies with reverse bias potential, the pinched resistance is voltage controlled and is nonlinear.

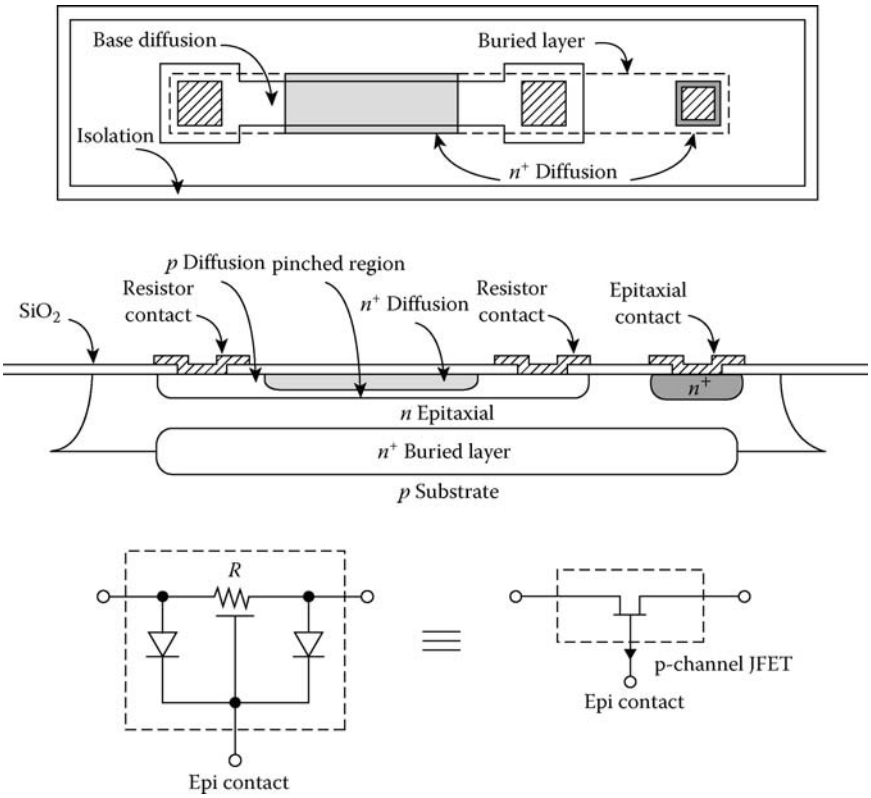


FIGURE 1.73 p-Type base-pinched resistor.

Absolute values for the base-pinched resistors can vary as much as ±50% due to large process variation in the fabrication of the active base region. The maximum voltage that can be applied across the base-pinched resistor of Figure 1.73 is restricted by the breakdown voltage between the n^+ diffusion layer and the p-type base diffusion. The breakdown voltage has a typical value around 6 V.

1.4.1.3 Epitaxial Resistors

Large values of sheet resistance can be obtained either by reducing the effective cross-sectional area of the resistor structure or by using a low doping concentration that forms the resistor body. The first technique is used to realize the pinched resistor while the second is used to realize the epitaxial resistor. Figure 1.74 shows an epitaxial resistor structure where the resistor is formed with a lightly doped epitaxial layer. For an epi thickness of 10 μm and a doping concentration of 10^{15} donor atoms/cm^3, this structure achieves a resistivity of 5 Ω-cm and an effective sheet resistance of 5 KΩ/N. The temperature coefficient of the epitaxial resistor is relatively high with typical values around +3000 ppm/°C. This large temperature variation is a direct consequence of the hole and electron mobilities undergoing more drastic variations against temperature at particularly low doping concentrations [13]. The maximum voltage that can be applied across the epitaxial resistor is significantly higher than that for the pinched resistor. This voltage

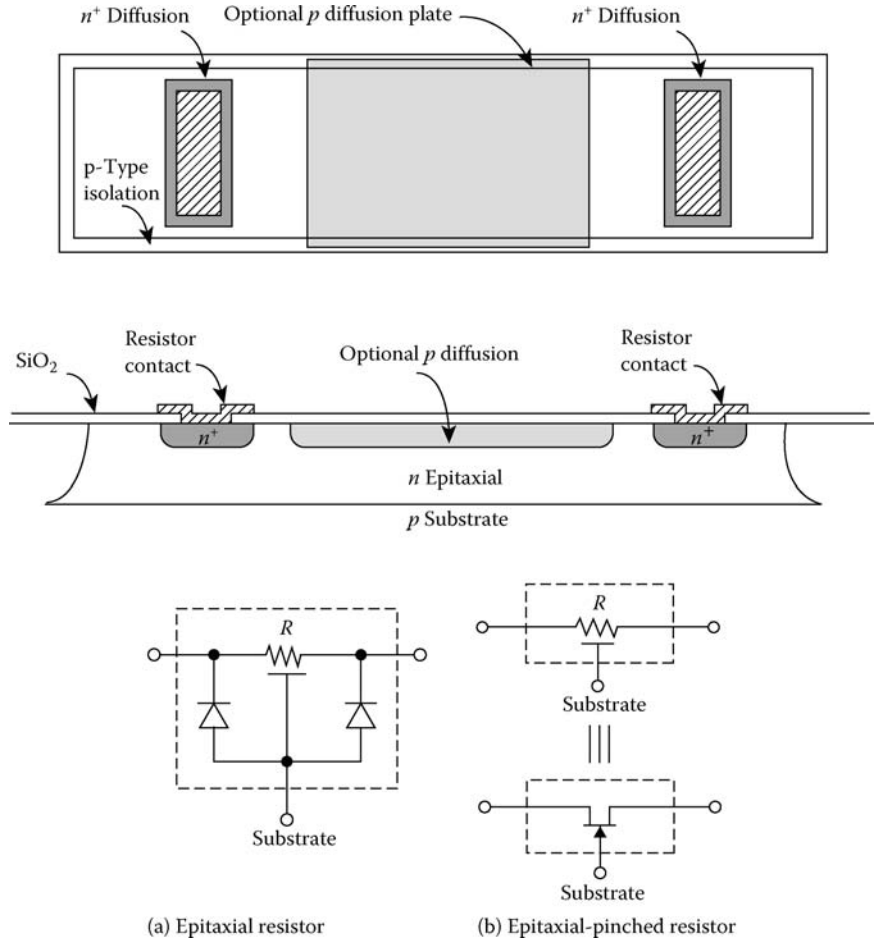


FIGURE 1.74 n-Type epitaxial and epitaxial-pinched resistors.

is set by the breakdown voltage between the n-type epi and the p-type substrate which varies inversely with the doping concentration of this pn junction.

Epitaxial-pinched resistors. By putting a p-type diffusion plate on top of the epitaxial resistor of Figure 1.74, even larger sheet resistance value can be obtained. The p-type diffusion plate overlaps the epi region and is electrically connected to the substrate through the p-type isolation. The epi layer is thus pinched between the p-type diffusion plate and the p-type substrate. When the n-type epi and the surrounding p-type regions is subject to a reverse bias potential, the junction depletion width extends into the epi region and effectively reduces the cross-sectional area. Typical sheet resistance values are between 4 and 5 KΩ/N. The epitaxial-pinched resistor behaves like an n-channel JFET, in which the effective channel width is controlled by the substrate voltage.

1.4.1.4 Ion-Implanted Resistors

Ion implantation is an alternative technique beside diffusion for inserting impurity atoms into a silicon wafer [17]. Commonly used impurities for implantation are the p-type boron atoms. The desired impurity atoms are first ionized and then accelerated to a high energy by an electric field. When a beam of these high-energy ions is directed at the wafer, the ions penetrate into exposed regions of the wafer surface. The penetration depth depends on the velocity at contact and is typically between 0.1 and 0.8 μm. The exposed regions of the wafer surface are defined by selectively etching a thick thermally grown SiO_2 layer that covers the wafer and functions as a barrier against the implanted ions. Unique characteristics of the ion-implantation technique include a precise control of the impurity concentration, uniformly implanted layers of impurity atoms, and no lateral diffusion. The structure of a p-type ion-implanted resistor is shown in Figure 1.75, where the p-type diffused regions at the contacts are used to achieve good ohmic contacts to the implanted resistor. The pn junction formed between the p-type implanted region and the n-type epitaxial layer must be reverse biased for electrical isolation. By connecting the epi region to a potential relatively more positive than the substrate potential, the conductive action due to the parasitic pnp transistor formed by the p-type implanted, the n-type epi, and the p-type substrate is also eliminated. Ion-implanted resistors exhibit relatively tight absolute value tolerance and excellent matching. Absolute value tolerance down to ±3% and matching tolerance of ±2% are typical performance.

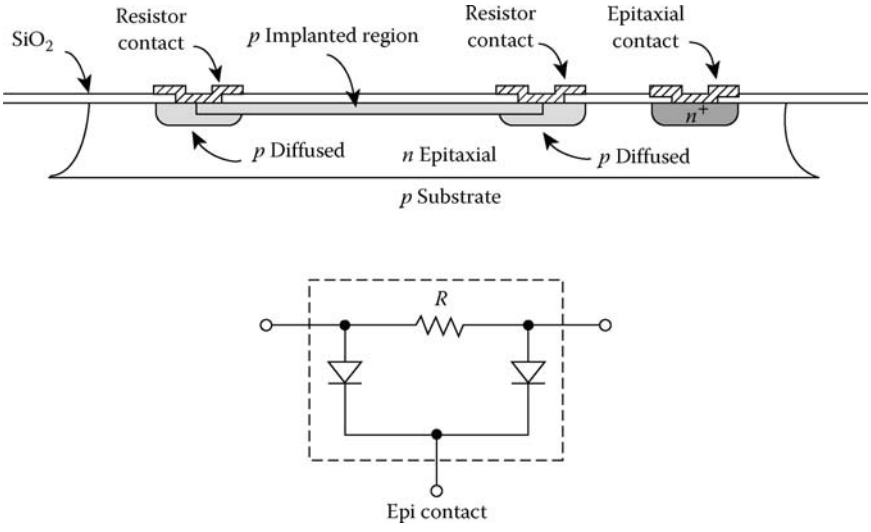


FIGURE 1.75 p-Type ion-implanted resistor.

TABLE 1.8 Typical Properties of Semiconductor Resistors

Resistor Type	Sheet ρ(Ω/N)	Absolute Tolerance (%)	Matching Tolerance (%)	Temperature Coefficient (ppm/°C)
Base-diffused	100–200	±20	±2 (5 μm wide)	+1500 to +2000
			±0.2 (50 μm wide)	—
Emitter-diffused	2–10	±20	±2	+600
Base-pinched	2–10 K	±50	±10	+2500
Epitaxial	2–5 K	±30	±5	+3000
Epitaxial-pinched	4–10 K	±50	±7	+3000
Ion-implanted	100–1000	±3	±2 (5 μm wide)	Controllable to ±100
			±0.15 (50 μm wide)	

Source: Gray, P.R. and Meyer, R.G., *Analysis and Design of Analog Integrated Circuits*, Wiley, New York, 1984, p. 119.

Table 1.8 provides a summary of the typical characteristics for the diffused, pinched, epitaxial, and ion-implanted resistors.

1.4.1.5 Thin-Film Resistors

Compared with diffused resistors, thin-film resistors offer advantages of a lower temperature coefficient, a smaller absolute value variation, and an excellent high-frequency characteristic. Commonly used resistive thin films are tantalum, Ni–Cr, Cr–SiO, and SnO_2. A typical thin-film resistor structure is shown in Figure 1.76, where a thin-film resistive layer is deposited on top of a thermally grown SiO_2 layer and a thin-film conductive metal layer is used to form the resistor contacts. The oxide layer functions as an insulating layer for the resistor. Various CVD techniques can be used to form the thin films [8]. The oxide passivation layer deposited on top of the resistive film and the conductive film protects the device surface from contamination. The electrical lumped model as shown in Figure 1.76 is adequate to characterize the high-frequency performance of the resistor. The parallel-plate capacitance formed

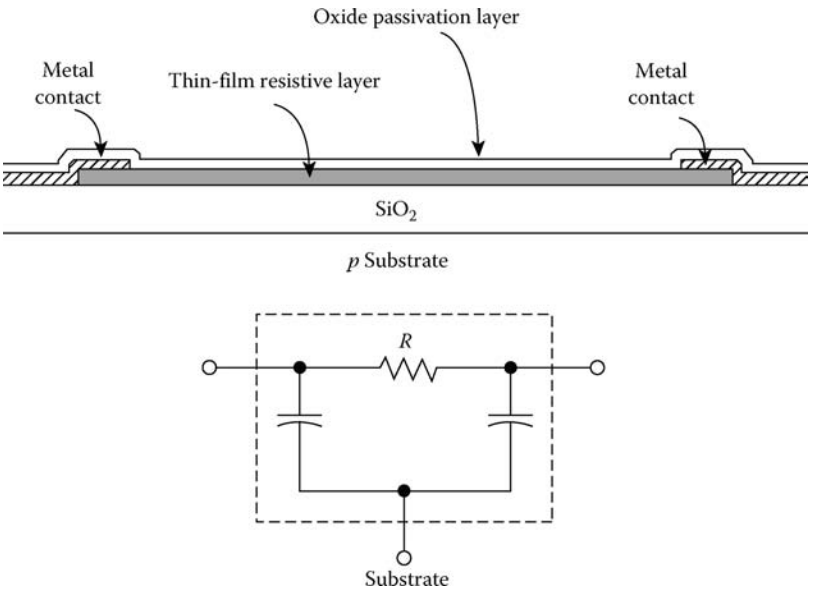


FIGURE 1.76 Thin-film resistor.

TABLE 1.9 Typical Characteristic of Thin-Film Resistors

Resistor Type	Sheet ρ(Ω/N)	Absolute Tolerance (%)	Matching Tolerance (%)	Temperature Coefficient (ppm/°C)
Ni–Cr	40–400	±5	±1	±100
Ta	10–1000	±5	±1	±100
SnO_2	80–4000	±8	±2	0–1500
Cr–SiO	30–2500	±10	±2	±50 to ±150

Source: Grebene, A.B., *Bipolar and MOS Analog Integrated Circuit Design*, Wiley, New York, 1984, p. 155.

between the thin-film resistive and the substrate is divided equally between the two capacitors. Table 1.9 provides a summary of the characteristics for some commonly used thin-film resistors.

1.4.2 Capacitors

Monolithic capacitors are widely used in analog and digital ICs for functions such as circuit stability, bandwidth enhancement, ac signal coupling, impedance matching, and charge storage cells. Capacitor structures available in monolithic form include pn junction, MOS, and polysilicon capacitors. pn junctions under reverse-biased conditions exhibit a nonlinear voltage-dependent capacitance. MOS and polysilicon capacitors, on the other hand, closely resemble the linear parallel-plate capacitor structure as shown in Figure 1.77. If the insulator thickness T of the parallel-plate structure is small compared with the plate width W and length L, the electric field between the plates is uniform (fringing field neglected). Under this condition the capacitance can be calculated by

$$C = \frac{\kappa\varepsilon_0}{T} WL \tag{1.245}$$

where κ is the relative dielectric constant of the insulating material and ε_0 is the permittivity constant in vacuum (8.854×10^{-14} F/cm).

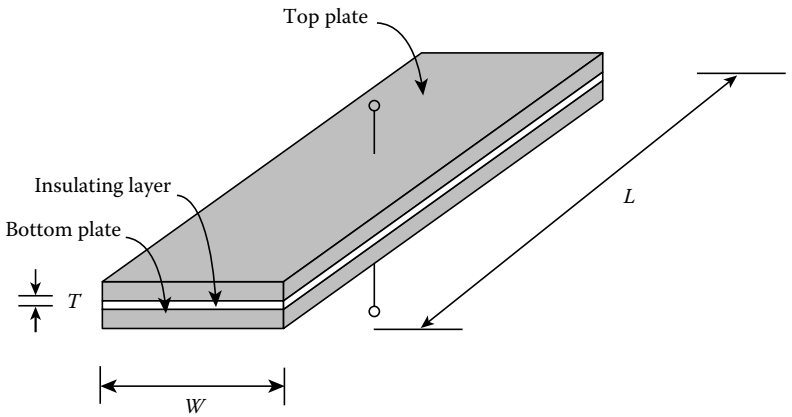


FIGURE 1.77 Structure of a parallel-plate capacitor.

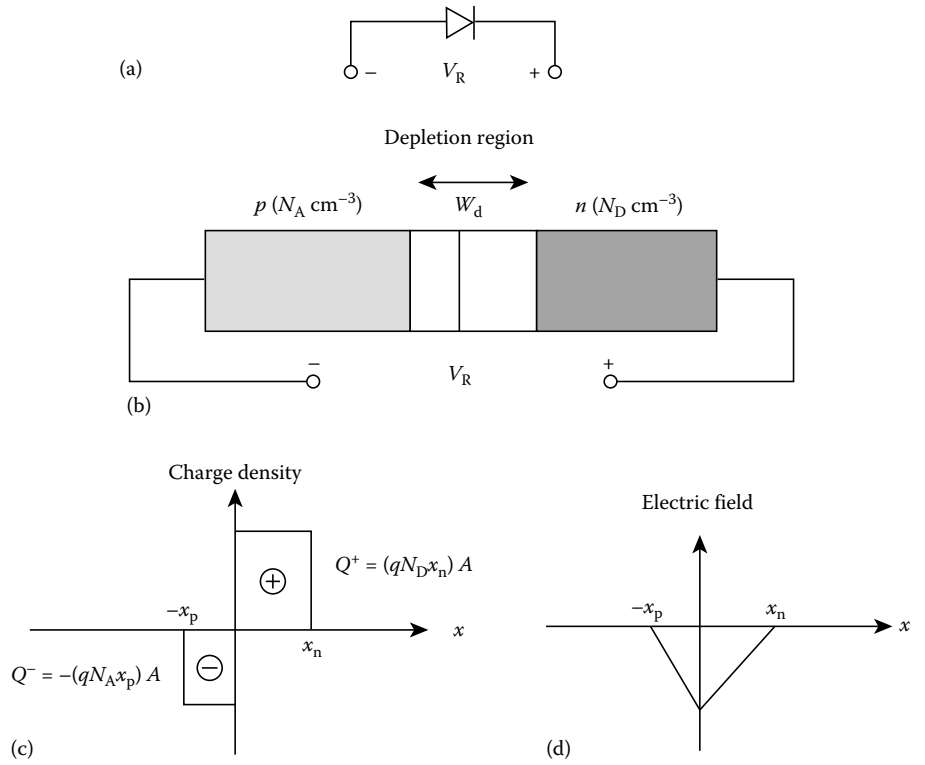


FIGURE 1.78 Abrupt p–n junction: (a) p–n junction symbol; (b) depletion region; (c) charge density within the depletion region; and (d) electric field.

1.4.2.1 Junction Capacitors

The structure of an *abrupt* pn junction is shown in Figure 1.78, where the doping is assumed uniform throughout the region on both sides. The acceptor impurity concentration of the p region is N_A atoms/cm^3 and the donor impurity concentration of the n region is N_D atoms/cm^3. When the two regions are brought in contact, mobile holes from the p region diffuse across the junction to the n region and mobile electrons diffuse from the n to the p region. This diffusion process creates a *depletion* region that is essentially free of mobile carriers (depletion approximation) and contains only fixed acceptor and donor ions. Ionized acceptor atoms are negatively charged and ionized donor atoms are positively charged. In equilibrium the diffusion process is balanced out by a drift process that arises from a *built-in* voltage ψ_o across the junction. This voltage is positive from the n region relative to the p region and is given by [17]

$$\psi_o = \frac{kT}{q} \ln \frac{N_A N_D}{n_i^2} \tag{1.246}$$

where

k is the Boltzmann constant (1.38×10^{-23} V · C/K)
T is the temperature in Kelvin (K)
q is the electron charge (1.60×10^{-19} C)
n_i(cm^{-3}) is the *intrinsic* carrier concentration in a pure semiconductor sample

For silicon at 300 K, $n_i \approx 1.5 \times 10^{10}$ cm^{-3}.

When the pn junction is subject to an applied reverse bias voltage V_R, the drift process is augmented by the external electric field and more mobile electrons and holes are pulled away from the junction. Because of this effect, the depletion width W_d and consequently the charge Q on each side of the junction vary with the applied voltage. A junction capacitor can thus be defined to correlate this charge–voltage relationship. The Poisson's equation relating the junction voltage $\phi(x)$ to the electric field $\xi(x)$ and the total charge Q is

$$\begin{aligned}\frac{d^2\varphi(x)}{dx^2} &= -\frac{d\xi(x)}{dx} = -\frac{q}{\varepsilon_S}(p - n + N_D - N_A) \\ &\approx \begin{cases} \dfrac{qN_A}{\varepsilon_S} & -x_p < x < 0 \\ -\dfrac{qN_D}{\varepsilon_S} & 0 < x < x_n \end{cases}\end{aligned} \tag{1.247}$$

where ε_S ($11.8\varepsilon_0 = 1.04 \times 10^{-12}$ F/cm) is the permittivity of the silicon material. The first integral of Equation 1.247 yields the electric field as

$$\xi(x) = \begin{cases} -\dfrac{qN_A}{\varepsilon_S}(x + x_p) & -x_p < x < 0 \\ -\dfrac{qN_D}{\varepsilon_S}(x_n + x) & 0 < x < x_n \end{cases} \tag{1.248}$$

The electric field is shown in Figure 1.78, where the maximum field strength occurs at the junction edge. This value is given by

$$|\xi_{max}| = \frac{qN_A}{\varepsilon_S}x_p = \frac{qN_D}{\varepsilon_S}x_n$$

The partial depletion width x_p on the p region and the partial depletion width x_n on the n region can then be related to the depletion width W_d as

$$\begin{aligned} x_p + x_n &= W_d \\ x_p &= \frac{N_D}{N_A + N_D}W_d \\ x_n &= \frac{N_A}{N_A + N_D}W_d \end{aligned}$$

Taking the second integral of Equation 1.247 yields the junction voltage

$$\varphi(x) = \begin{cases} \dfrac{qN_A}{\varepsilon_S}\left(\dfrac{x_p^2}{2} + x_p x + \dfrac{x^2}{2}\right) & -x_p < x < 0 \\ \dfrac{qN_D}{\varepsilon_S}\left(\dfrac{x_n x_p}{2} + x_n x - \dfrac{x^2}{2}\right) & 0 < x < x_n \end{cases} \tag{1.249}$$

where the voltage at x_p is arbitrarily assigned to be zero. The total voltage $\psi_0 + V_R$ can be expressed as

$$\psi_o + V_R = \varphi(x_n) = \frac{qN_D}{2\varepsilon_S}\left(1 + \frac{N_D}{N_A}\right)x_n^2$$

Finally, the depletion width W_d and the total charge Q in terms of the total voltage across the junction can be derived to be

$$W_d = \left[\frac{2\varepsilon_S}{q}(\psi_o + V_R)\left(\frac{1}{N_A} + \frac{1}{N_D}\right)\right]^{1/2}$$
$$|Q| = A(qN_A x_p) = A(qN_D x_n) = A\left[2q\varepsilon_S(\psi_o + V_R)\left(\frac{1}{N_A} + \frac{1}{N_D}\right)^{-1}\right]^{1/2} \tag{1.250}$$

The junction capacitance is thus

$$C_j = \left|\frac{dQ}{dV_R}\right| = A\left[\frac{q\varepsilon_S}{2}\left(\frac{1}{\psi_o + V_R}\right)\left(\frac{1}{N_A} + \frac{1}{N_D}\right)^{-1}\right]^{1/2}$$
$$= \frac{C_{jo}}{\left(1 + \frac{V_R}{\psi_o}\right)^{1/2}} \tag{1.251}$$

where

A is the effective cross-sectional junction area
C_{jo} is the value of C_j for $V_R = 0$

If the doping concentration in one side of the pn junction is much higher than that in the other, the depletion width and the junction capacitance can be simplified to

$$W_d = \left[\frac{2\varepsilon_S}{qN_L}(\psi_o + V_R)\right]^{1/2} \tag{1.252}$$

$$C_j = A\left[\frac{\varepsilon_S qN_L}{2}\left(\frac{1}{\psi_o + V_R}\right)\right]^{1/2} \tag{1.253}$$

where N_L is the concentration of the lightly doped side. Figure 1.79 displays the junction capacitance per unit area as a function of the total voltage $\psi_o + V_R$ and the concentration on the lightly doped side of the junction [3].

In silicon bipolar technology the base–emitter, the base–collector, and the collector–substrate junctions under reverse bias are often utilized for realizing a junction capacitance. The collector–substrate junction has only a limited use since it can only function as a shunt capacitor due to the substrate being connected to an ac ground.

Base–collector junction capacitor. A typical base–collector capacitor structure is shown in Figure 1.80 together with an equivalent lumped circuit model. A heavily doped n^+ buried layer is used to minimize the series resistance R_C. For the base–collector junction to operate in reverse bias, the n-type collector must be connected to a voltage relatively higher than the voltage at the p-type base. The junction breakdown voltage is determined by BV_{CBO} of the npn transistor, which has a typical value between 25 and 50 V.

Base–emitter junction capacitor. Figure 1.81 shows a typical base–emitter capacitor structure where the parasitic junctions D_{BC} and D_{SC} must always be in reverse bias. The base–emitter junction achieves the highest capacitance per unit area among the base–collector, base–emitter, and collector–substrate junctions due to the relatively higher doping concentrations in the base and emitter regions.

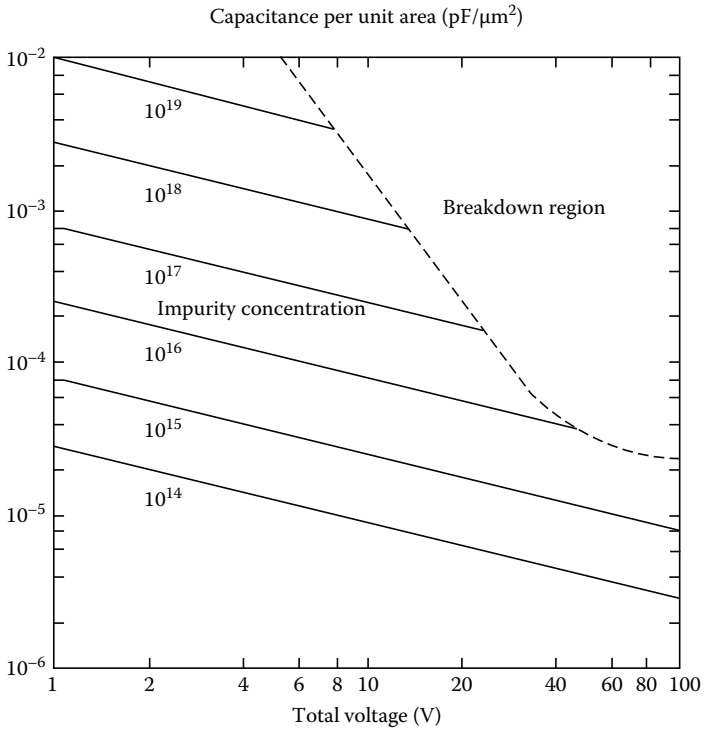


FIGURE 1.79 Junction capacitance as a function of the total voltage and the concentration on the lightly doped side.

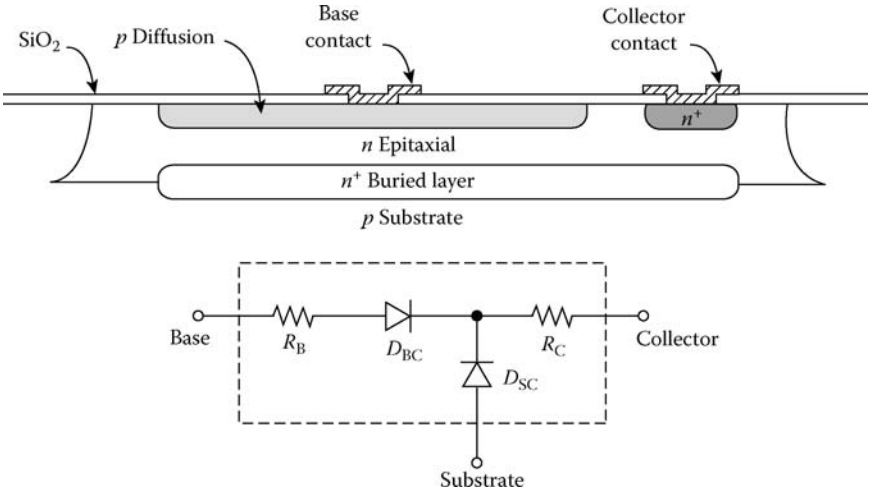


FIGURE 1.80 Base–collector junction capacitor.

For the base–emitter junction to operate in reverse bias, the n-type emitter must be connected to a voltage relatively higher than the voltage at the p-type base. The breakdown voltage of the base–emitter junction is relatively low, determined by the BV_{EBO} of the npn transistor, which has a typical value of about 6 V.

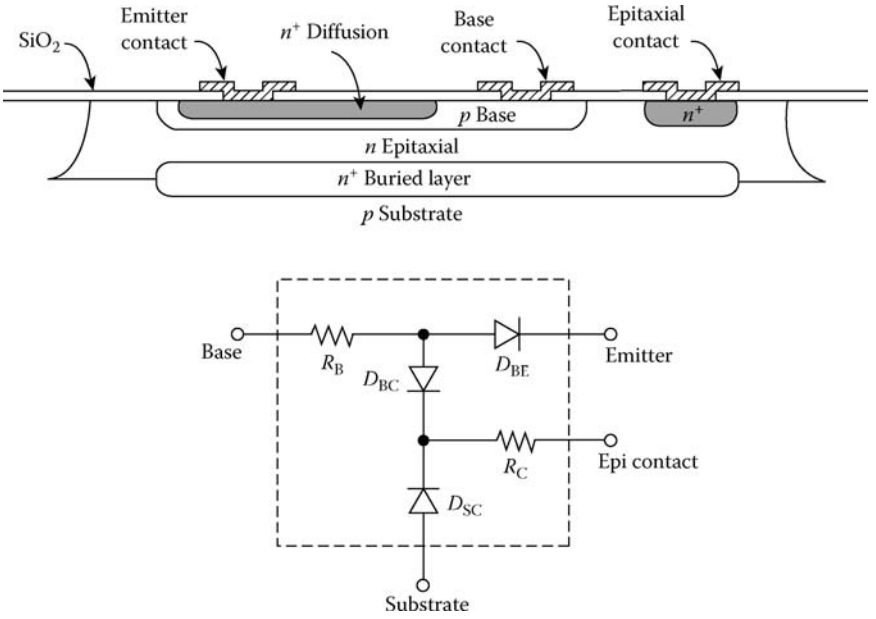


FIGURE 1.81 Base–emitter junction capacitor.

1.4.2.2 MOS Capacitors

MOS capacitors are preferable and commonly used in ICs since they are linear and not confined to a reverse-biased operating condition as in the junction capacitors. The structure of a MOS capacitor is shown in Figure 1.82, where by means of a local oxidation process a thin oxide layer is thermally grown on top of a heavily doped n^+ diffusion layer. The oxide layer has a typical thickness between 500 and 1500 Å (Å $= 10^{-10}$ m $= 10^{-4}$ μm) and functions as the insulating layer of the parallel-plate capacitor. The top plate is formed by overlapping the thin oxide area with a deposited layer of conductive metal. The bottom-plate diffusion layer is heavily doped for two reasons: to minimize the bottom-plate resistance and to minimize the depletion width at the oxide-semiconductor interface when the capacitor operates in the *depletion* and *inversion* modes [17]. By keeping the depletion width small, the effective capacitance is dominated by the parallel-plate oxide capacitance. The MOS capacitance is thus given by

$$C = \frac{\kappa_{ox}\varepsilon_0}{T} A \tag{1.254}$$

where

κ_{ox} is the relative dielectric constant of SiO_2 (2.7 to 4.2)
ε_0 is the permittivity constant
T is the oxide thickness
A is the area defined by the thin oxide layer

In practice, a thin layer of silicon nitride (Si_3N_4) is often deposited on the thin oxide layer and is used to minimize the charges inadvertently introduced in the oxide layer during oxidation and subsequent processing steps. These *oxide charges* are trapped within the oxide and can cause detrimental effect to the capacitor characteristic [17]. The silicon nitride assimilates an additional insulating layer and effectively

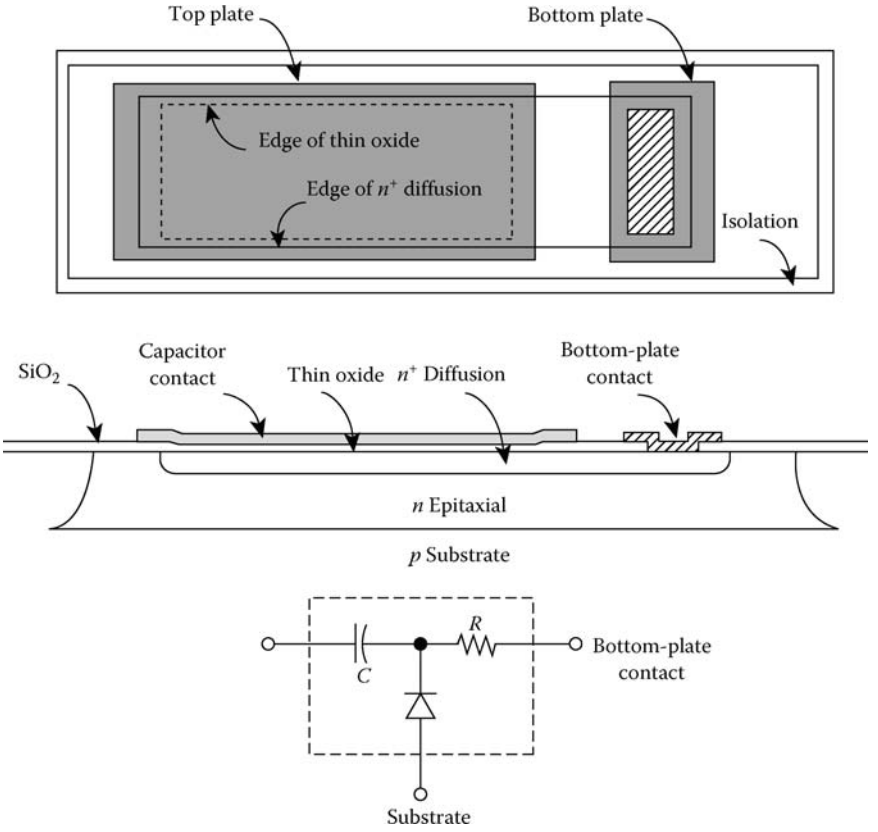


FIGURE 1.82 MOS capacitor.

creates an additional capacitor in series with the oxide capacitor. The capacitance for such a structure can be determined by an application of *Gauss's law*. It is given by

$$C = \frac{\varepsilon_o}{\left(\frac{T_{ni}}{\kappa_{ni}}\right) + \left(\frac{T_{ox}}{\kappa_{ox}}\right)} A \tag{1.255}$$

where

T_{ni} and T_{ox} are the thickness of the silicon nitride and oxide layers, respectively

κ_{ni} (2.7 to 4.2) and κ_{ni} (3.5 to 9) are the relative dielectric constant of oxide and silicon nitride, respectively

In the equivalent circuit model of Figure 1.82, the parasitic junction between the p-type substrate and the n-type bottom plate must always be reverse biased. The bottom-plate contact must be connected to a voltage relatively higher than the substrate voltage.

1.4.2.3 Polysilicon Capacitors

Polysilicon capacitors are conveniently available in MOSFET technology, where the gate of the MOSFET transistor is made of polysilicon material. Polysilicon capacitors also assimilate the parallel-plate capacitor. Figure 1.83 shows a typical structure of a polysilicon capacitor, where a thin oxide is deposited on top of a polysilicon layer and serves as an insulating layer between the top-plate metal layer and the

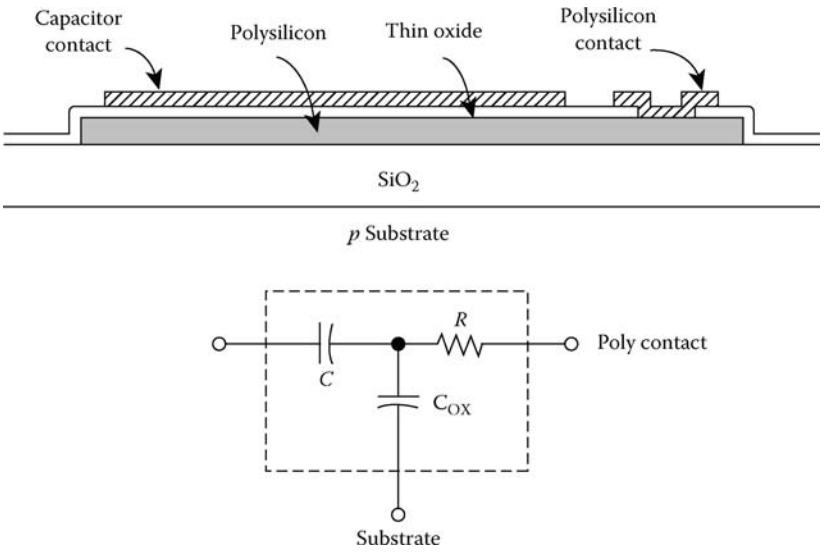


FIGURE 1.83 Polysilicon capacitor.

bottom-plate polysilicon layer. The polysilicon region is isolated from the substrate by a thick oxide layer that forms a parasitic parallel-plate capacitance between the polysilicon layer and the substrate. This parasitic capacitance must be accounted for in the equivalent circuit model. The capacitance of the polysilicon capacitor is determined by either Equation 1.254 or 1.255 depending on whether a thin silicon nitride is used in conjunction with the thin oxide.

1.4.3 Inductors

Planar inductors have been implemented using a variety of substrates such as standard PC boards, ceramic and sapphire hybrids, monolithic GaAs [24], and more recently monolithic silicon [18]. In the early development of silicon technology, planar inductors were investigated [26], but the prevailing lithographic limitations and relatively large inductance requirements (for low-frequency applications) resulted in excessive silicon area and poor performance. Reflected losses from the conductive silicon substrate were a major contribution to low inductor *Q*. Recent advances in silicon IC processing technology have achieved fabrication of metal width and metal spacing in the low micrometer range and thus allow many more inductor turns per unit area. Also, modern oxide-isolated processes with multilayer metal options allow thick oxides to help isolate the inductor from the silicon substrate. Practical applications of monolithic inductors in low-noise amplifiers, impedance matching amplifiers, filters and microwave oscillators in silicon technologies have been successfully demonstrated [19,20].

Monolithic inductors are especially useful in high-frequency applications where inductors of a few nano-Henrys of inductance are sufficient. Inductor structures in monolithic form include strip, loop, and spiral inductors. Rectangular and circular spiral inductors are by far the most commonly used structures.

1.4.3.1 Rectangular Spiral Inductors

The structure of a rectangular spiral inductor is shown in Figure 1.84, where the spiral loops are formed with the top metal layer M_2 and the connector bridge is formed with the bottom metal layer M_1. Using the top metal layer to form the spiral loops has the advantage of minimizing the parasitic metal-to-substrate

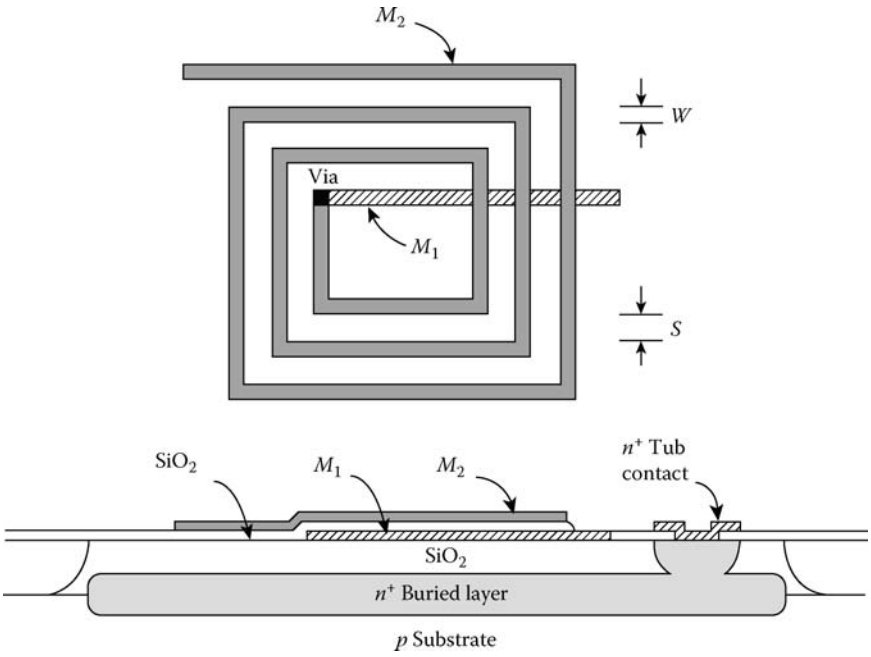


FIGURE 1.84 Rectangular spiral inductor.

capacitance. The metal width is denoted by W and the metal spacing is denoted by S. The total inductance is given by

$$L_T = \sum_{i=1}^{4N} L_S(i) + 2 \cdot \sum_{i=1}^{4N-1} \sum_{j=i+1}^{4N} L_M(ij) \tag{1.256}$$

where

N is the number of turns
$L_S(i)$ is the *self inductance* of the rectangular metal segment i
$L_M(ij)$ is the *mutual inductance* between metal segments i and j

The self-inductance is due to the magnetic flux surrounding each metal segment. The mutual inductance is due to the magnetic flux coupling around every two parallel metal segments and has a positive value if the currents applied to the metal conductors flow in the same direction and a negative value otherwise. Perpendicular metal segments have negligible mutual inductance.

The self-inductance and mutual inductance for straight rectangular conductors can be determined by the *geometric mean distance* method [10], in which the conductors are replaced by equivalent straight filaments whose basic inductive characteristics are well known.

Self-inductance. The self-inductance for the rectangular conductor of Figure 1.85 depends on the conductor length L, the conductor width W, and the conductor thickness T. The static self-inductance is given by [9,10].

$$L_S = 2L\left[\ln\left(\frac{2L}{GMD}\right) - 1.25 + \left(\frac{AMD}{L}\right) + \left(\frac{\mu_r}{4}\right)\zeta\right](nH) \tag{1.257}$$

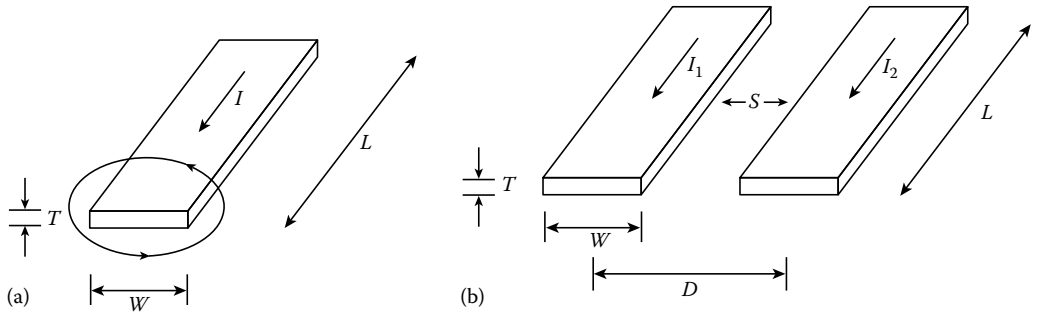


FIGURE 1.85 Calculation of (a) self-inductance and (b) mutual inductance for parallel rectangular conductors.

where

μ_r is the relative permeability constant of the conductor
GMD is the geometric mean distance
AMD is the arithmetic mean distance
ζ is a frequency-dependent parameter that equals 1 for direct and low-frequency alternating currents and approaches 0 for very high-frequency alternating currents

The AMD and GMD for the rectangular conductor of Figure 1.85 are

$$\text{AMD} = \left(\frac{W+T}{3}\right)$$
$$\text{GMD} = \begin{cases} 0.22313 \cdot (W+T) & T \to 0 \\ 0.22360 \cdot (W+T) & T = \text{W}/2 \\ 0.223525 \cdot (W+T) & T \to \text{W} \end{cases} \tag{1.258}$$

The rectangular dimensions L, W, and T are normalized to the centimeter in the preceding expressions.

Mutual inductance. The mutual inductance for the two parallel rectangular conductors of Figure 1.85 depends on the conductor length L, the conductor width W, and the conductor thickness T, and the distance D separating the conductor centers. The static mutual inductance is [10]

$$L_{\text{M}} = 2L\alpha(\text{nH}) \tag{1.259}$$

where

$$\alpha = \ln\left[\left(\frac{L}{\text{GMD}}\right) + \left[1 + \left(\frac{L}{\text{GMD}}\right)^2\right]^{1/2}\right] - \left[1 + \left(\frac{\text{GMD}}{L}\right)^2\right]^{1/2} + \left(\frac{\text{GMD}}{L}\right)$$

and

$$\text{GMD} = \exp(\ln D - \beta) \tag{1.260}$$

$$\beta = \begin{cases} \frac{1}{12}\left(\frac{D}{W}\right)^{-2} + \frac{1}{60}\left(\frac{D}{W}\right)^{-4} + \frac{1}{168}\left(\frac{D}{W}\right)^{-6} + \frac{1}{360}\left(\frac{D}{W}\right)^{-8} \\ + \frac{1}{660}\left(\frac{D}{W}\right)^{-10} + \cdots \\ 0.1137 \quad \text{for} \quad D = W \end{cases}$$

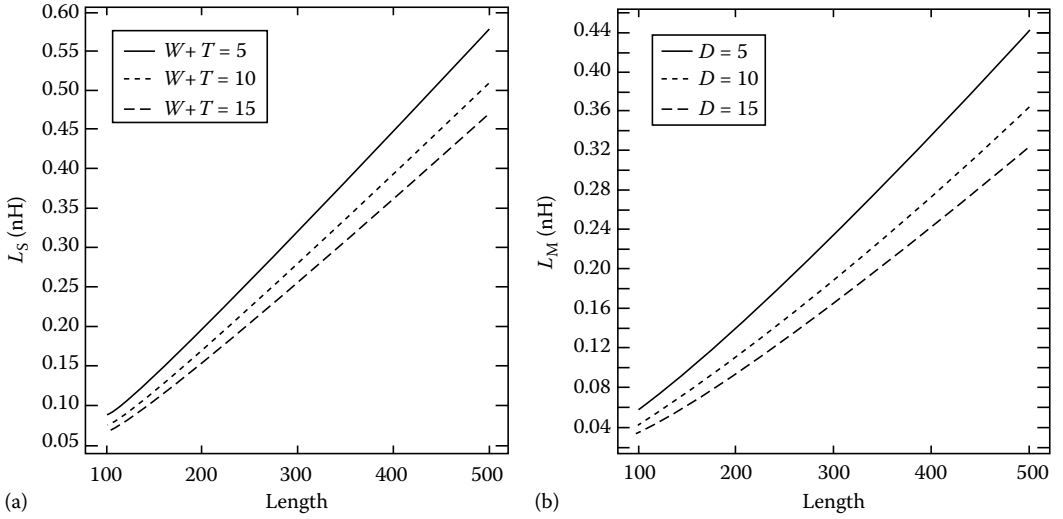


FIGURE 1.86 (a) Self-inductance as a function of width, thickness, and length for rectangular conductors. (b) Mutual inductance as a function of distance and length for rectangular conductors ($W = 5$, $T = 0$).

The GMD closed-form expression Equation 1.260 is valid for rectangular conductors with small thickness-to-width ratios T/W. As the thickness T approaches the width W, the GMD approaches the distance D and the GMD is no longer represented by the above closed-form expression. Figure 1.86 shows plots of the self inductance and the mutual inductance as expressed in Equations 1.257 and 1.259, respectively. The conductor dimensions are given in μm (μm $= 10^{-4}$ cm).

For the inductor structure of Figure 1.84 it is important to emphasize that since the spiral loops are of nonmagnetic metal material, the total inductance depends only on the geometry of the conductors and not on the current strength. At high-frequencies, especially those above the *self-resonant* frequency of the inductor, the *skin effect* due to current crowding toward the surface and the propagation delay as the current traverses the spiral must be fully accounted for [16,22]. The ground-plane effect due to the inductor image must also be considered regardless of the operation frequency.

An equivalent lumped model for the rectangular spiral inductor of Figure 1.84 is shown in Figure 1.87. This model consists of the total inductance L_T, the accumulated metal resistance R_S, the coupling

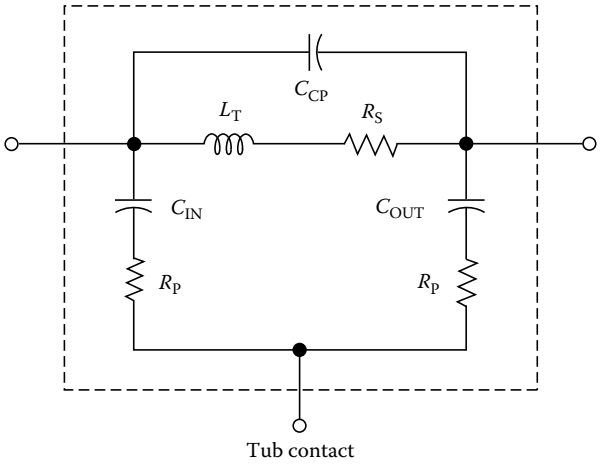


FIGURE 1.87 Electrical model for the spiral inductor.

capacitance C_{CP} between metal segments due to the electric fields in both the oxide region and the air region, the parasitic capacitances C_{IN} and C_{OUT} from the metal layers to the buried layer [2,11,15], and the buried-layer resistance R_P. Since the spiral structure of Figure 1.84 is not symmetrical, the parasitic capacitors C_{IN} and C_{OUT} are not the same, though the difference is relatively small. The self-resonant frequency can be approximated using the circuit model of Figure 1.87 with one side of the inductor being grounded. For simplicity, let $C_{IN} = C_{OUT} + C_P$ and neglect the relatively small coupling capacitor C_{CP}, the self-resonant frequency is given by

$$f_R = \frac{1}{2\pi} \frac{1}{\sqrt{L_T C_P}} \left[\frac{1 - R_S^2 \left(\frac{C_P}{L_T} \right)}{1 - R_P^2 \left(\frac{C_P}{L_T} \right)} \right]^{1/2} \tag{1.261}$$

Transformer structures. Transformers are often used in high-performance analog ICs that require conversions between single-ended signals and differential signals. In monolithic technology, transformers can be fabricated using the basic structure of the rectangular spiral inductor. Figure 1.88 shows a planar interdigitated spiral transformer that requires only two metal layers M_1 and M_2. The structure of Figure 1.89, on the other hand, requires three layers of metal for which the top metal layer M_3 is used for the upper spiral, the middle metal layer M_2 is used for the lower spiral, and the bottom metal layer M_1 is used for the two connector bridges. This structure can achieve a higher inductance per unit than that of Figure 1.88 due to a stronger magnetic coupling between the upper spiral and the lower spiral through a relatively thin oxide layer separating metal layers M_2 and M_3. An equivalent lumped model is

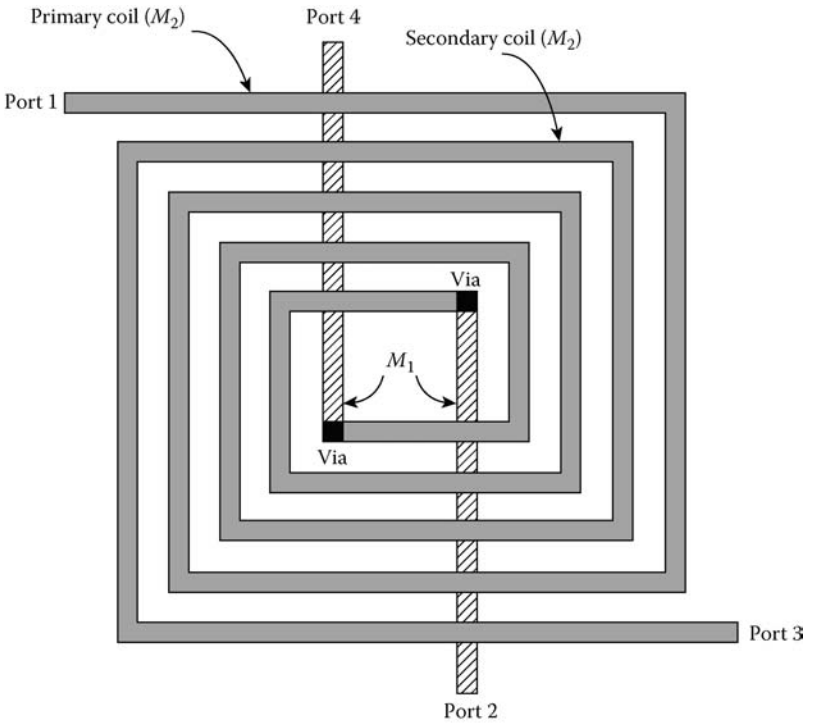


FIGURE 1.88 Rectangular spiral transformer I.

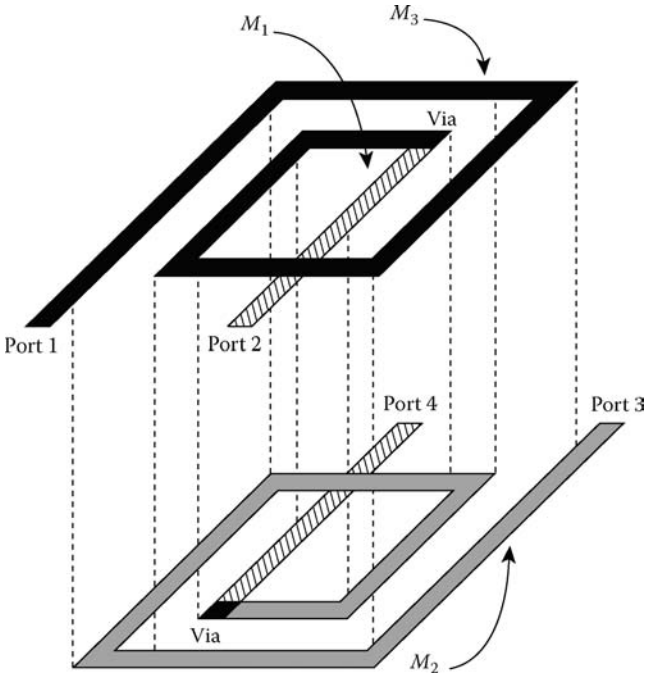


FIGURE 1.89 Rectangular spiral transformer II.

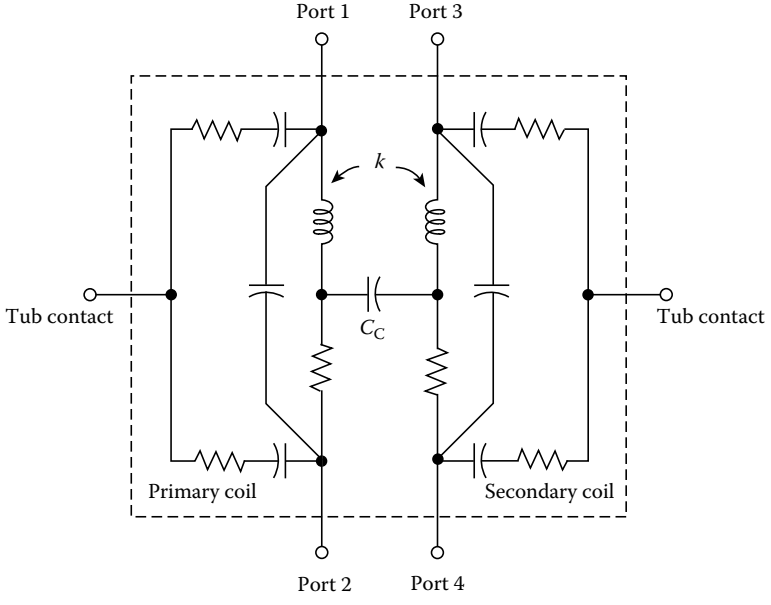


FIGURE 1.90 Electrical model for the spiral transformer.

shown in Figure 1.90. In addition to all the circuit elements of the two individual spiral inductors, there are also a magnetic coupling factor k and a coupling capacitance C_C between the primary and secondary coils.

1.4.3.2 Circular Spiral Inductors

The structure of a concentric circular spiral inductor is shown in Figure 1.91 where the circular loops share the same center point. The top metal layer M_2 is used for the circular conductors and the bottom metal layer M_1 is used for the connector bridge. The metal width is denoted by W and the spacing between two adjacent loops is denoted by S. The total inductance is given by

$$L_T = \sum_{i=1}^{N} L_S(i) + 2 \cdot \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} L_M(ij) \tag{1.262}$$

where

N is the number of circular turns
$L_S(i)$ is the *self-inductance* of the circular conductor i
$L_M(ij)$ is the *mutual inductance* between conductors i and j

Self-inductance. Consider the single circular conductor of Figure 1.92a that has a radius R and a width W. A current I applied to this conductor produces a magnetic flux encircled by the loop and another magnetic flux inside the conductor itself. The inductance associated with the former and the latter magnetic flux component is referred to as the *external* self-inductance and the *internal* self-inductance, respectively. The external self-inductance characterizing the change in the encircled magnetic flux to the change in current is [25].

$$L_S = \mu(2R - \delta)\left[\left(1 - \frac{k^2}{2}\right)K(k) - E(k)\right](\text{nH}) \tag{1.263}$$

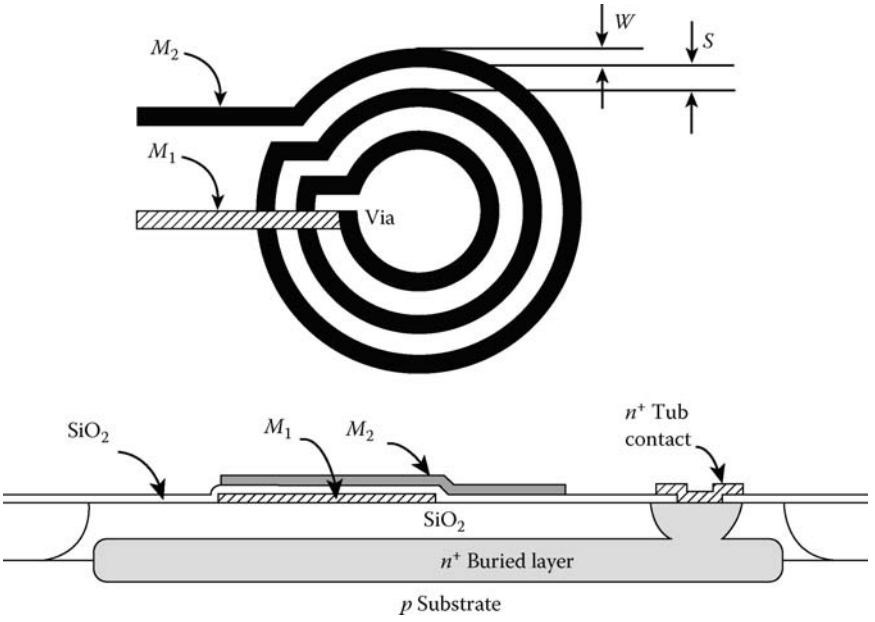


FIGURE 1.91 Concentric circular spiral inductor.

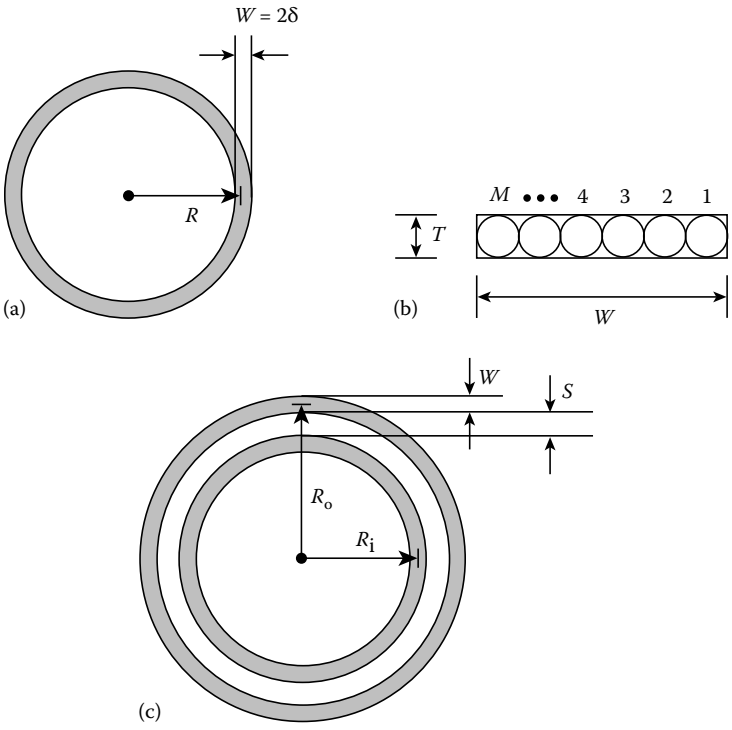


FIGURE 1.92 Calculation of self-inductance and mutual inductance for circular conductors. (a) External self-inductance; (b) internal self-inductance; and (c) mutual inductance.

where

$$k^2 = \frac{4R(R-\delta)}{(2R-\delta)^2} \tag{1.264}$$

and μ is the permeability of the conductor (equals 4π nH/cm for nonmagnetic conductors), and δ is one-half the conductor width W. $K(k)$ and $E(k)$ are the *complete elliptic integrals* of the first and second kind, respectively, and are given by

$$K(k) = \int_0^{\pi/2} \frac{d\phi}{\sqrt{1-k^2\sin^2\phi}}$$

$$E(k) = \int_0^{\pi/2} \sqrt{1-k^2\sin^2\phi}\, d\phi$$

The internal self-inductance is determined based on the concept of magnetic field energy. As shown in Figure 1.92b, the flat circular conductor is first approximated by an M number of round circular conductors [14] that are electrically in parallel and each conductor has a diameter equal to the thickness T of the flat conductor. The internal self-inductance of each round conductor is then determined as [25].

$$L = \frac{\mu}{8\pi}(\text{nH/cm})$$

The internal self-inductance of the flat conductor thus equals the parallel combination of these M components

$$L_S \approx \frac{\mu}{4}\left\{\sum_{i=1}^{M}[R - \delta + T(i - 0.5)]^{-1}\right\}^{-1}(\text{nH}) \tag{1.265}$$

where $R - \delta + T(i - 0.5)$ is the effective radius from the center of the loop to the center of the round conductor i. The typical contribution from the internal self-inductance of Equation 1.265 is less than 5% the contribution from the external self-inductance of Equation 1.263.

Mutual inductance. The mutual inductance of the two circular loops of Figure 1.92c depends on the inner radius R_i and the outer radius R_o. For any two adjacent loops of the circular spiral inductor, the outer radius is related to the inner radius by the simple relation $R_o = R_i + (W + S)$. The mutual inductance is determined based on the *Neumann's* line integral given as follows:

$$L_M = \frac{\mu}{4\pi}\int_C\int_C \frac{d\mathbf{l}_1 \cdot d\mathbf{l}_2}{D}$$

where

$d\mathbf{l}_1 \cdot d\mathbf{l}_2$ represents the dot product of the differential lengths
D is the distance separating the differential $\mathbf{l}_1$ vector and $\mathbf{l}_2$ vector

The static mutual inductance [25] is

$$L_M = \mu\sqrt{R_i R_o}\left[\left(\frac{2}{k} - k\right)K(k) - \frac{2}{k}E(k)\right](\text{nH}) \tag{1.266}$$

where

$$k^2 = \frac{4R_i R_o}{(R_i + R_o)^2} \tag{1.267}$$

Figure 1.93 shows plots of the external self-inductance and the mutual inductance as expressed in Equations 1.263 and 1.266, respectively. The conductor dimensions are given in μm.

As in the rectangular spiral inductor, the ground-plane effect and the retardation effect of the circular spiral inductor must be fully accounted for. The circuit model of Figure 1.87 can be used to characterize the electrical behavior of the circular inductor.

A comparison between the rectangular spiral of Figure 1.84 and the circular spiral of Figure 1.91 is shown in Figure 1.94, where the total inductance L_T is plotted against the turn number N. Both inductors have the same innermost dimension, the same conductor width, space, and thickness. The dimensions are given in μm, and the ground-plane effect and the retardation effect are not considered. For a given turn number, the rectangular spiral yields a higher inductance per semiconductor area than the circular spiral. Figure 1.95 shows a plot of the inductor Q vs. the total inductance of the same spiral inductors under consideration. Due to a higher inductance per length ratio, the Q of the circular inductor is higher than that of the rectangular inductor, about 10% for high inductance values.

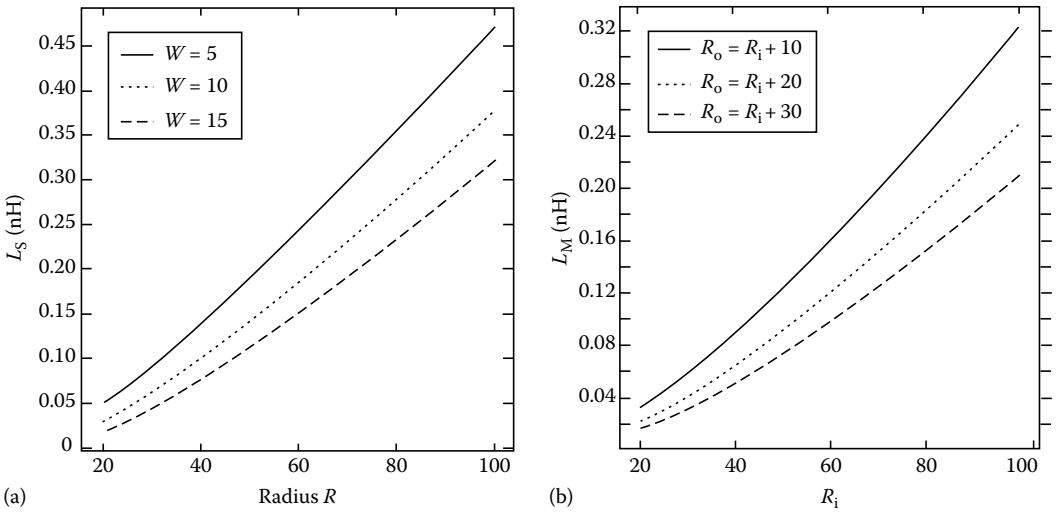


FIGURE 1.93 (a) External self-inductance as a function of radius and width for circular conductors. (b) Mutual inductance as a function of radii R_i and R_o for circular conductors.

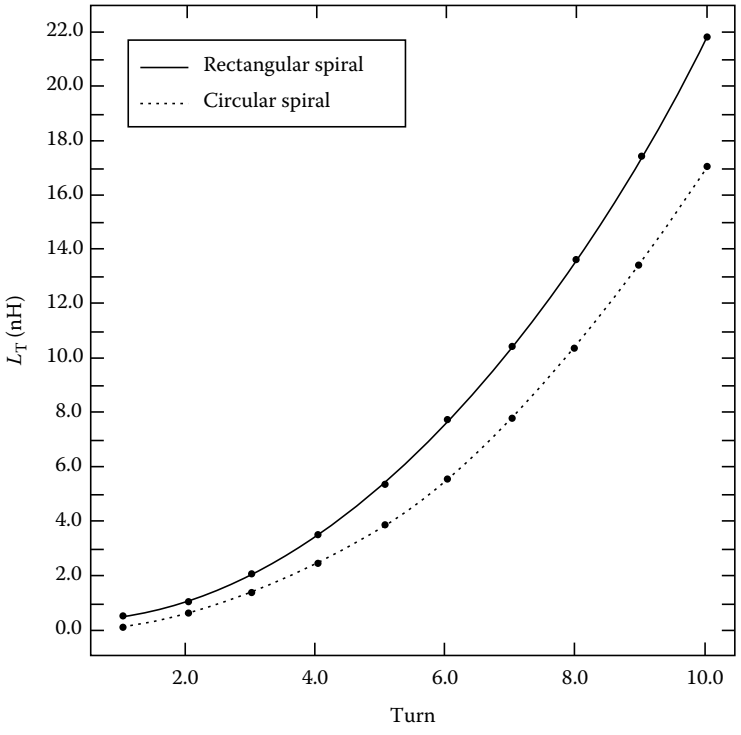


FIGURE 1.94 Total static inductance vs. turn number for the rectangular and circular inductors. The ground-plane effect is neglected. Innermost center dimension is 88 by 88, $W = 6$, $S = 3$, $T = 1.2$.

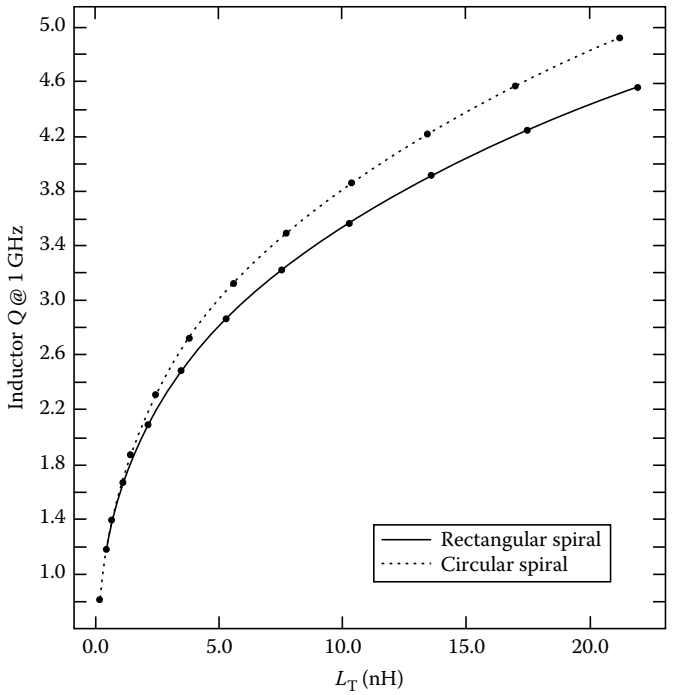


FIGURE 1.95 Inductor Q vs. total inductance for the rectangular and circular inductors. Metal sheet resistance is 25 mΩ/N. Innermost dimension is 88, $W = 6$, $S = 3$, $T = 1.2$.

References

1. I. Bahl and P. Bhartia, *Microwave Solid State Circuit Design*, New York: Wiley, 1988.
2. T. G. Bryant and J. A. Weiss, Parameters of microstrip transmission lines and of coupled pairs of microstrip lines, *IEEE Trans. Microwave Theory Tech.*, MTT-16, 1021–1027, 1968.
3. H. R. Camenzind, *Electronic Integrated Systems Design*, New York: Van Nostrand Reinhold, 1972.
4. E. M. Conwell, Properties of silicon and germanium, *Proc. IRE*, 46, 1281–1300, 1958.
5. R. Garg and I. J. Bahl, Characteristics of coupled microstrip lines, *IEEE Trans. Microwave Theory Tech.*, MTT-27, 700–705, 1979.
6. F. R. Gleason, Thin-film microelectronic inductors, in *Proc. Nat. Electron. Conf.*, 1964, pp. 197–198.
7. P. R. Gray and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 2nd ed., New York: Wiley, 1984.
8. A. B. Grebene, *Bipolar and MOS Analog Integrated Circuit Design*, New York: Wiley, 1984.
9. H. M. Greenhouse, Design of planar rectangular microelectronic inductor, *IEEE Trans. Parts, Hybrids, Packaging*, PHP-10, 101–109, 1974.
10. F. W. Grover, *Inductance Calculations*, New York: Van Nostrand, 1946.
11. E. Hammerstad and O. Jensen, Accurate models for microstrip computer-aided design, *IEEE MTT-S Dig.*, 80, 407–409, 1980.
12. J. C. Irwin, Resistivity of bulk silicon and of diffused layers in silicon, *Bell Syst. Tech. J.*, 41, 387–410, 1962.
13. C. Jacoboni, C. Canali, G. Ottaviani, and A. A. Quaranta, A review of some charge transport properties of silicon, *Solid State Electron.*, 20, 77–89, 1977.

14. R. L. Kemke and G. A. Burdick, Spiral inductors for hybrid and microwave applications, in *Proc. Electron. Components Conf.*, 1974, pp. 152–161.
15. M. Kirschning and R. H. Jansen, Accurate wide-range design equations for the frequency-dependent characteristics of parallel-coupling microstrip lines, *IEEE Trans. Microwave Theory Tech.*, MTT-32, 83–90, 1984.
16. D. Krafcsik and D. Dawson, A close-form expression for representing the distributed nature of the spiral inductor, *IEEE MTT-S Dig.*, 86, 87–92, 1986.
17. R. S. Muller and T. I. Kamins, *Device Electronics for Integrated Circuits*, 2nd ed., New York: Wiley, 1986.
18. N. M. Nguyen and R. G. Meyer, Si IC-compatible inductors and LC passive filters, *IEEE J. Solid-State Circuits*, 25, 1028–1031, 1990.
19. N. M. Nguyen and R. G. Meyer, A Si bipolar monolithic RF bandpass amplifier, *IEEE J. Solid-State Circuits*, 27, 123–127, 1992.
20. N. M. Nguyen and R. G. Meyer, A 1.8-GHz monolithic LC voltage-controlled oscillator, *IEEE J. Solid-State Circuits*, 27, 444–450, 1992.
21. N. M. Nguyen and R. G. Meyer, Start-up and frequency stability in high-frequency oscillators, *IEEE J. Solid-State Circuits*, 27, 810–820, 1992.
22. M. Parisot, Y. Archambault, D. Pavlidis, and J. Magarshack, Highly accurate design of spiral inductors for MMIC's with small size and high cut-off frequency characteristics, *IEEE MTT-S Dig.*, 84, 106–110, 1984.
23. E. Pettenpaul, H. Kapusta, A. Weisgerber, H. Mampe, J. Luginsland, and I. Wolff, CAD models of lumped elements on GaAs up to 18 GHz, *IEEE Trans. Microwave Theory Tech.*, 36, 294–304, 1988.
24. R. A. Pucel, Design considerations for monolithic microwave circuits, *IEEE Trans. Microwave Theory Tech.*, MTT-29, 513–534, 1981.
25. S. Ramon, J. R. Whinnery, and T. V. Duzer, *Fields and Waves in Communication Electronics*, 2nd ed., New York: Wiley, 1984.
26. R. M. Warner, Jr., and J. N. Fordemwalt, *Integrated Circuits*, New York: McGraw-Hill, 1965.

1.5 Chip Parasitics in Analog Integrated Circuits

Martin A. Brooke

The parasitic elements in electronic devices and interconnect limit the performance of all ICs. No amount of improvement in device performance or circuit design can completely eliminate these effects. Thus, as circuit speeds increase, unaccounted for interconnect parasitics become a more and more common cause of analog IC design failure. Hence, the causes, characterization, and modeling of significant interconnect parasitics are essential knowledge for good analog IC design [1–4].

1.5.1 Interconnect Parasitics

The parasitics due to the wiring used to connect devices together on chip produce a host of problems. Unanticipated feedback through parasitic capacitances can cause unwanted oscillation. Mismatch due to differences in interconnect resistance contribute to unwanted offset voltages. For very-high-speed ICs, the inductance of interconnects is both a useful tool and a potential cause of yield problems.

Even the interactions between interconnect lines are both important and very difficult to model. So too are the distributed interactions of resistance, capacitance, and (in high-speed circuits) inductance that produce transmission line effects.

1.5.1.1 Parasitic Capacitance

Distributed capacitance of IC lines is perhaps the most important of all IC parasitics. It can lower the bandwidth of amplifiers, alter the frequency response of filters, and cause oscillations.

Physics. Every piece of IC interconnect has capacitance to the substrate. In the case of silicon circuitry, the substrate is conductive and connected to an ac ground, thus there is a capacitance to ground from every circuit node due to the interconnect. Figure 1.96 illustrates this substrate capacitance interconnect parasitic. The capacitance value will depend on the total area of the interconnect, and on the length of edge associated with the interconnect. This edge effect is due to the nonuniformity of the electric field at the interconnect edges. The nonuniformity of the electric field at edges is such that the capacitance value is larger for a given area of interconnect near the edge than elsewhere.

In addition to the substrate capacitance, all adjacent pieces of an interconnect will have capacitance between them. This capacitance is classified into two forms, overlap capacitance, and parallel line capacitance (also known as proximity capacitance). Overlap capacitance occurs when two pieces of interconnect cross each other, while parallel line capacitance occurs when two interconnect traces run close to each other for some distance.

When two lines cross each other, the properties of the overlapping region will determine that size of the overlap capacitance. The electric field through a cross section of two overlapping lines is illustrated in Figure 1.97. The electric field becomes nonuniform near the edges of the overlapping region, producing an edge-dependent capacitance term. The capacitance per unit area at the edge is always greater than elsewhere and, if the overlapping regions are small, the edge capacitance effect can be significant.

The size of parallel line capacitance depends on the distance for which the two lines run side by side and on the separation of the lines. Since parallel line capacitance occurs only at the edges of an interconnect, the electric field that produces it is very nonuniform. This particular nonuniformity, as illustrated in Figure 1.98, makes the capacitance much smaller for a given area of interconnect than either overlap or substrate capacitance. Thus, two lines must run parallel for some distance for this capacitance to be important. The nonuniformity of the electric field makes the dependence of the capacitance on line separation highly nonlinear, as a result the capacitance value decreases much more rapidly with separation than it would if it depended linearly on the line separation.

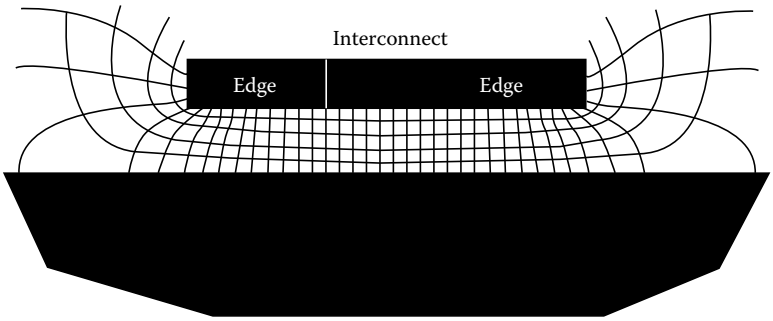


FIGURE 1.96 Substrate capacitance. The electric field distorts at the edges, making the capacitance larger there than elsewhere.

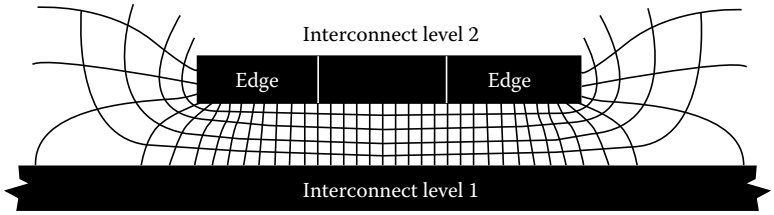


FIGURE 1.97 Overlap capacitance. The bottom interconnect level will have edges into and out of the page with distorted electric field similar to that shown for the top level of interconnect.

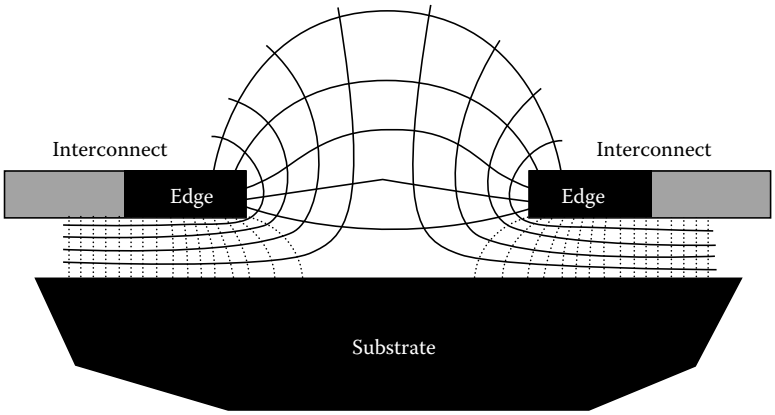


FIGURE 1.98 Parallel line capacitance. Only the solid field lines actually produce line-to-line capacitance, the dashed lines form substrate capacitance.

Modeling. In the absence of significant interconnect resistance effects, all of the parasitic capacitances can be modeled with enough accuracy for most analog circuit design applications by dissecting the interconnect into pieces with similar capacitance characteristics and adding up the capacitance of each piece to obtain a single capacitance term. For example, the dissected view of a piece of interconnect with substrate capacitance is shown in Figure 1.99. The interconnect has been dissected into squares that fall into three

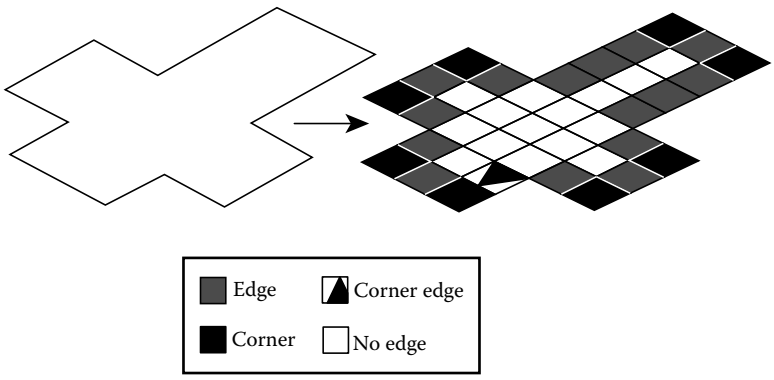


FIGURE 1.99 Determining substrate capacitance. The capacitance of each square in the dissected interconnect segment is summed.

classes: two types of edges, and one center type. The capacitance to the substrate for each of these squares in parallel and thus the total capacitance of the interconnect segment is simply the sum of the capacitance of each square. If the substrate capacitance contribution of each square has been previously measured or calculated, the calculation of the total interconnect segment substrate capacitance involves summing each type of squares capacitance multiplied by the number of squares of that type in the segment.

The accuracy of this modeling technique depends solely on the accuracy of the models used for each type of square. For example, in Figure 1.99, the accuracy could be improved by adding one more type of edge square to those that are modeled. One of these squares has been shaded differently in the figure and is called the corner edge square.

For the nonedge pieces of the interconnect the capacitance is approximately a parallel-plate capacitance and can be computed from Equation 1.268.

$$C = \frac{A \cdot \varepsilon_r \cdot \varepsilon_0}{t} \tag{1.268}$$

where

A is the area of the square or piece of interconnect
t is the thickness of the insulation layer beneath the interconnect
ε_r is the relative dielectric constant of the insulation material
ε_0 is the dielectric constant of free space

For silicon ICs insulated with silicon dioxide the parameters are given in Table 1.10.

The capacitance of edge interconnect pieces will always be larger than nonedge pieces. The amount by which the edge capacitance increases will depend on the ratio of the size of the piece of interconnect and the thickness of the insulation layer beneath the interconnect. If the interconnect width is significantly larger than the thickness of the insulation then edge effects are probably small and can be ignored. However, when thin lines are used in ICs the edge effects are usually significant. The factor by which the edge capacitance can increase over the parallel-plate approximation can easily be as high as 1.5 for thin lines.

The modeling of overlap capacitance is handled in the same fashion as substrate capacitance. The region where interconnect lines overlap is dissected into edges and nonedges and the value of capacitance for each type of square summed up to give a total capacitance between the two circuit nodes associated with each piece of interconnect that overlaps. The area of overlap between the two layers of interconnect can be used as A in Equation 1.268, while that separation between the layers can be used as t. The strong distortion of the electric fields will increase the actual value above this idealized computed value by a factor that depends on the thickness of the lines. This factor can be as high as 2 for thin lines.

Parallel line capacitance can also be handled in a manner similar to that used for substrate and overlap capacitance. However, we must now locate *pairs* of edge squares, one from each of the adjacent interconnect lines. In Figure 1.100, one possible pairing of the squares from adjacent pieces of interconnect is shown. The capacitance for each type of pair of squares is added together, weighted by the number of pairs of each type to get a single capacitance that connects the circuit nodes associated with each interconnect line.

The effect on the capacitance of the spacing between pairs must be either measured or computed for each possible spacing, and type of pair of squares. One approach to this is to use a table of measured or

TABLE 1.10 Parameters for Calculation of Substrate Capacitance in Silicon ICs Insulated with Silicon Dioxide

Parameter	Value
ε_r	3.9
ε_0	$8.854 \cdot 10^{-12}$ F/m
t	$1–5 \cdot 10^{-6}$ m

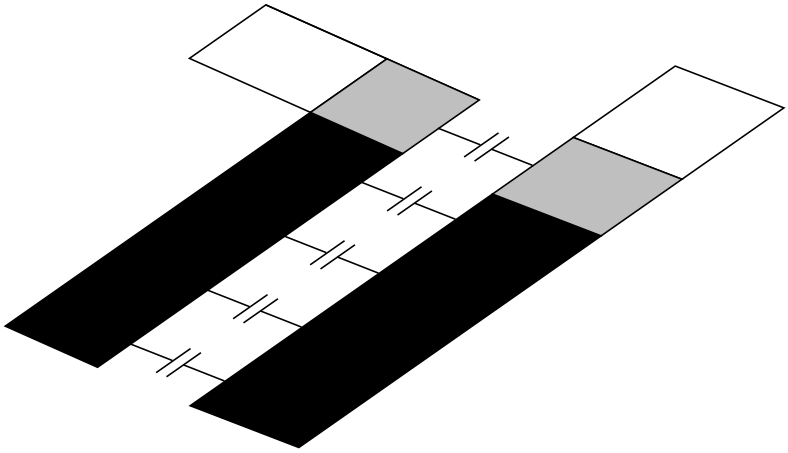


FIGURE 1.100 Determining parallel line capacitance. The differently shaded pairs of squares are different types and will each have a different capacitance between them.

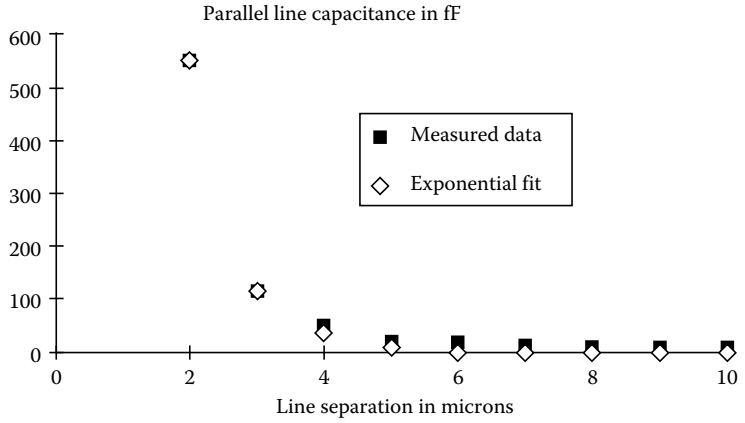


FIGURE 1.101 Parallel line capacitance measured from a silicon IC. The diamonds are an exponential fit to the data (using Equation 1.269). The fit is excellent at short separations when the capacitance is largest.

computed capacitances and separation distances. The measured parallel line capacitance between silicon IC lines for a variety of separations is presented in Figure 1.101. From the figure, we see that the capacitance value decreases exponentially with line separation. Thus an exponential fit to measured or simulated data is good choice for computing the capacitance [7,8].

Equation 1.269 can be used to predict the parallel line capacitance C for each type of pair of edge squares. L is the length of the edge of the squares, and the parameters C_c and S_d are computed or fit to measured coupling capacitance data like that in Figure 1.101.

$$C = C_c \cdot L \cdot e^{-(s/S_d)} \tag{1.269}$$

Effects on circuits. The effects that parasitic capacitances are likely to produce in circuits range from parametric variations, such as reduced bandwidth, to catastrophic failures, such as amplifier oscillation. Each type of parasitic capacitance produces a characteristic set of problems. Being aware of these typical problems will ease diagnosis of actual, or potential, parasitic capacitance problems.

Substrate capacitance usually causes lower than expected bandwidth in amplifiers and lowering of the poles in filters. The capacitance is always to ac ground and thus increases device and circuit capacitances to ground. Thus, circuit nodes that have a dominant effect on amplifier bandwidth, or filter poles, should be designed to have as little substrate capacitance as possible. Another, more subtle, parametric variation that can be caused by substrate capacitance is frequency-dependent mismatch. For example, if the parasitic capacitance to ground is different between the two inputs of a differential amplifier, then, for fast transient signals, the amplifier will appear unbalanced. This could limit the accuracy high-speed comparators, and is sometimes difficult to diagnose since the error only occurs at high speeds.

Overlap and parallel line capacitance can cause unwanted ac connections to be added to a circuit. These connections will produce crosstalk effects and can result in unstable amplifiers. The output interconnect and input interconnect of high-gain or high-frequency amplifiers must thus be kept far apart at all times. Care must be taken to watch for series capacitances of this type. For example, if the output and input interconnect of an amplifier both cross the power supply interconnect, unwanted feedback can result if the power supply line is not well ac grounded. This is a very common cause of IC amplifier oscillation. Because of the potential for crosstalk between parallel or crossing lines, great care should also be taken to keep weak (high-impedance) signal lines away from strong (low-impedance) signal lines.

1.5.1.2 Parasitic Resistance

For analog IC designers, the second most important interconnect parasitic is resistance. This unexpected resistance can cause both parametric problems, such as increased offset voltages, and catastrophic problems such as amplifier oscillation (for example, poorly sized power supply lines can cause resistive positive feedback paths in high gain amplifiers called ''ground loops''). To make matters worse, the resistivity of IC interconnect has been steadily increasing as the line widths of circuits have decreased.

Physics. Except for superconductors, all conductors have resistance. A length of interconnect used in an IC is no exception. The resistance of a straight section of interconnect is easily found by obtaining the resistance per square for the particular interconnect layer concerned, and then adding up the resistance of each of the series of squares that makes up the section. This procedure is illustrated in Figure 1.102.

For more complicated interconnect shapes the problem of determining the resistance between two points in the interconnect is also more complex. The simplest approach is to cut the interconnect up into rectangles and assume each rectangle has a resistance equal to the resistance per square of the interconnect material times the number of full and partial squares that will fit along the direction of current flow in the rectangle [5]. This scheme works whenever the direction of current flow is clear; however, for corners and intersections of interconnect the current flow is in fact quite complex. Figure 1.103 shows the kind of current flow that can occur in an interconnect section with complex geometry.

Modeling. To account for the effects of complex current flows the resistance of complex interconnect geometries must be determined by measurement or simulation. One simple empirical approach is to cut out sections of resistive material in the same shape as the interconnect shape to be modeled, and then

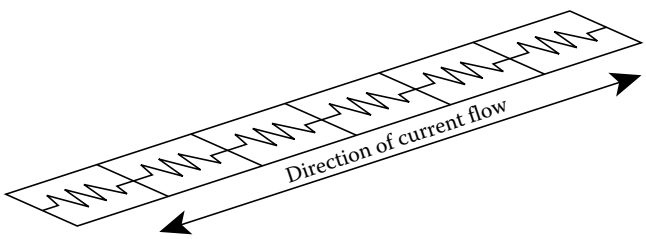


FIGURE 1.102 Determining the resistance of a length of interconnect. Each square has the same resistance regardless of size.

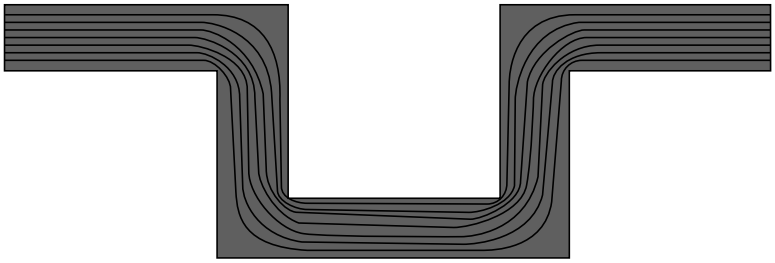


FIGURE 1.103 Current flow in a complex interconnect geometry.

measure the resistance. The resistance for other materials can be found by multiplying by the ratio of the respective resistances per square of the two materials.

Once the resistance has been found for a particular geometry it can be used for any linear scaling of that geometry. For most types of IC interconnect all complex geometries can be broken up into relatively few important subgeometries. If tables of the resistance of these subgeometries for various dimensions and connection patterns are obtained, the resistance of quite complex shapes can be accurately calculated by connecting the resistance of each subgeometry together and calculating the resistance of the connected resistances. This calculation can usually be performed quickly by replacing series and parallel connected resistor pairs with their equivalents. The process of breaking a complex geometry into subgeometries, constructing the equivalent connected resistance, and forming a single resistance for an interconnect section is illustrated in Figure 1.104.

Effects on circuits. The resistance of interconnect can have both parametric and catastrophic effects on circuit performance. Even small differences in the resistance on either input side of a differential amplifier can lead to increased offset voltage. Thus, when designing differential circuits care must be taken to make the interconnect identical on both input sides, as this ensures that the same resistance is present in both circuits.

The resistance of power supply interconnect can lead to both parametric errors in the voltages supplied and catastrophic failure due to oscillation. If power supply voltages are assumed to be identical in two parts of a circuit and, due to interconnect resistance, there is a voltage drop from one point to the next, designs that rely on the voltages being the same may fail. In high-gain and feedback circuits the resistance of the ground and power supply lines may become an unintentional positive feedback resistance which could lead to oscillation. Thus output and input stages for high-gain amplifiers will usually require separate ground and power supply interconnects. This ensures that no parasitic resistance is in a feedback path.

When using resistors provided in an IC process, the extra resistance provided by the interconnect may cause inaccuracies in resistor values. This would be most critical for small resistance values. The only solution in this case is to accurately compute the interconnect resistance. Since most resistance layers

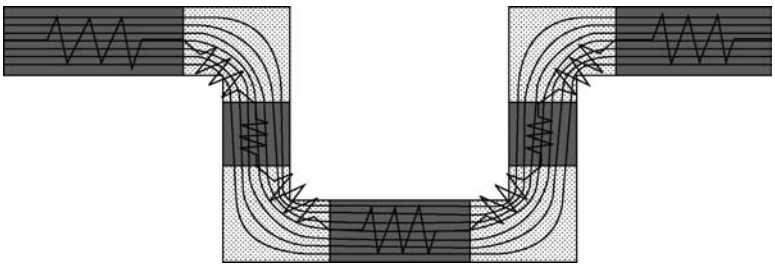


FIGURE 1.104 The process of breaking a complex geometry into subgeometries, constructing the equivalent connected resistance, and forming a single resistance for an interconnect section. In this example, only two subgeometries are used: a corner subgeometry and a basic rectangular subgeometry.

provided in analog IC processes are just a form of high resistivity interconnect, the methods described here for accurately computing the resistance of interconnect are also useful for predicting the resistance of resistors to be fabricated.

1.5.1.3 Parasitic Inductance

In high-speed ICs the inductance of long lines of interconnect becomes significant. In IC technologies that have an insulating substrate, such as gallium arsenide (GaAs) and silicon on insulator (SOI), reasonably high-performance inductive devices can be made from interconnect. In technologies with conductive substrates, resistive losses in the substrate restrict the application of interconnect inductance. High-frequency circuits are often tuned using interconnect inductance and capacitance (*LC*) to form a narrow bandpass filter or tank circuit, and *LC* transmission lines, or stubs, made from interconnect are useful for impedance matching. There is much similarity between this use of parasitic inductance and the design of microstripline-printed circuit boards. The major difference being that inductance does not become significant in IC interconnect until frequencies in the gigahertz are reached.

In order to make a good interconnect inductance, there are two requirements. First, there must not be any resistive material within range of the magnetic field of the inductance. If this occurs then induced currents flowing in the resistive material will make the inductor have high series resistance (low *Q* factor). This would make narrow bandwidth bandpass filters difficult to make using the inductance, and make transmission lines made from the interconnect lossy. The solution is to have an insulating substrate, or to remove the substrate from beneath the inductor.

The second requirement for large inductance is to form a coil or other device to concentrate the magnetic field lines. Within the confines of current IC manufacturing, spiral inductors, like that illustrated in Figure 1.105 are the most common method used to obtain useful inductances.

1.5.1.4 Transmission Line Behavior

Two types of transmission line behavior are important in ICs, *RC* transmission lines and *LC*/*RLC* transmission lines. For gigahertz operation inductive transmission lines are important. These can be lossy *RLC* transmission lines if a conductive substrate such as silicon is used, or nearly lossless *LC* transmission lines if an insulating substrate such as GaAs is used. The design of inductive transmission lines is very similar to designing microstripline-printed circuit boards. At lower frequencies of

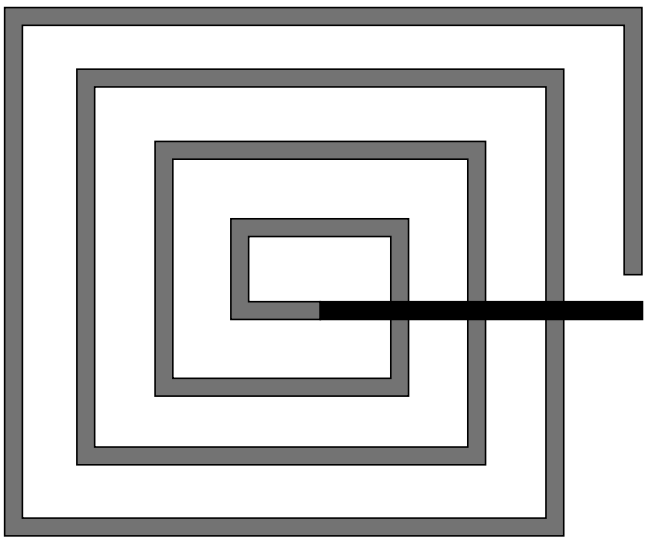


FIGURE 1.105 Spiral inductance used in insulated substrate ICs for gigahertz frequency operation.

10–1000 MHz resistive capacitive (*RC*) transmission lines are important for long low resistivity interconnect lines or short high resistivity lines.

RC transmission lines are of concern to analog circuit designers working in silicon ICs. When used correctly, an interconnect can behave as though it were purely capacitive in nature. However, when a higher resistivity interconnect layer, such as polysilicon or diffusion is used, the distributed resistance and capacitance can start to produce transmission line effects at relatively short distances. Similarly, for very long signal distribution lines or power supply lines, if they are not correctly sized, transmission line behavior ensues.

Physics. One method for modeling distributed transmission line interconnect effects is lumped equivalent modeling [6]. This method is useful for obtaining approximate models of complex geometries quickly, and is the basis of accurate numerical finite element simulation techniques. For analog circuit designers the conversion of interconnect layout sections into lumped equivalent models also provides an intuitive tool to understanding distributed transmission line interconnect behavior.

To be able to model a length of interconnect as a lumped *RC* equivalent, the error between the impedance of the interconnect when correctly treated as a transmission line, and when replaced with the lumped equivalent, must be kept low. If this error is *e*, then it can be shown that the maximum length of interconnect that can be modeled as a simple *RC T* or Π network is given in Equation 1.270. In the equation, *R* is the resistance per square of the particular type of interconnect used, *C* is the capacitance per unit area, and ω is the frequency of operation in radians per second.

$$D < \sqrt{\frac{3 \cdot e}{\omega \cdot R \cdot C}} \tag{1.270}$$

This length can be quite short. Consider the case of a polysilicon interconnect line in a 1.2 μm CMOS process that has a resistance per square of 40 Ω a capacitance per unit are of 0.1 fF/μm^2. For an error *e* of 10% the maximum line length of minimum width line that can be treated as a lumped T or Π network for various frequencies is given in Table 1.11. Longer interconnect lines than this must be cut up into lengths less than or equal to the length given by Equation 1.270.

Modeling. The accurate modeling of distributed transmission line effects in ICs is best performed with lumped equivalent circuits. These circuits can be accurately extracted by dissecting the interconnect geometry into lengths that are, at most, as long as the length given by Equation 1.270. These lengths are then modeled by either a *T* or Π lumped equivalent *RC* network. The extraction of the resistance and capacitance for these short interconnect sections can now follow the same procedures as were described in Sections 1.5.1.2 and 1.5.1.1. The resulting *RC* network is then an accurate transmission line model of the interconnect. Figure 1.106 shows an example of this process.

Effects on circuits. Several parametric and catastrophic problems can arise due to unmodeled transmission line behavior. Signal propagation delays in transmission lines are longer than predicted by a single lumped capacitance and resistance model of interconnect. Thus, ignoring the effects of transmission lines can result in slower circuits than expected. If the design of resistors for feedback networks

TABLE 1.11 The Maximum Length of Minimum Width Polysilicon Line That Can Be Modeled with a Single Lumped *RC T* or π Network and Remain 10% Accurate

Frequency (MHz)	Length (μm)
10	1262
100	399
1000	126

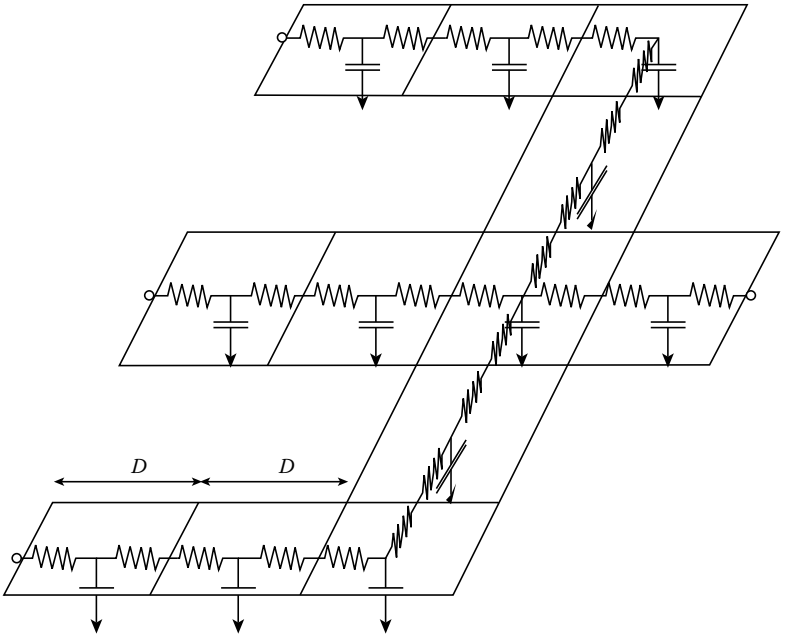


FIGURE 1.106 The extraction of an accurate *RC* transmission line model for resistive interconnect. The maximum allowable length D is computed from Equation 1.270.

results in long lengths of the resistive interconnect used to make the resistors, these resistors may in fact be *RC* transmission lines. The extra delay produced by the transmission line may well cause oscillation of the feedback loops using these resistors. The need for decoupling capacitors in digital and analog circuit power supplies is due to the *RC* transmission line behavior of the power supply interconnect. Correct modeling of the *RC* properties of the power distribution interconnect is needed to see whether fast power supply current surges will cause serious changes in the supply voltage or not.

1.5.1.5 Nonlinear Interconnect Parasitics

A number of types of interconnect can have nonlinear parasitics. These nonlinear effects are a challenge to model accurately because the effect can change with the operating conditions of the circuit. A conservative approach is to model the effects as constant at the worst likely value they can attain. This is adequate for predicting parameters, like circuit bandwidth, that need only exceed a specification value. If the specifications call for accurate prediction of parasitics then large nonlinear parasitics are generally undesirable and should be avoided.

Most nonlinear interconnect parasitics are associated with depletion or inversion of the semiconductor substrate. A diffusion interconnect is insulated from conducting substrates such as silicon by a reversed biased diode. This diode's depletion region width varies with the interconnect voltage and results in a voltage-dependent capacitance to the substrate. For example, the diffusion interconnect in Figure 1.107 has voltage-dependent capacitance to the substrate due to a depletion region. The capacitance value depends on the depletion region thickness, which depends on the voltage difference between the interconnect and the substrate.

$$C = C_0 \cdot \left(1 - \left(\frac{V_s}{\varphi_B}\right)\right)^M \tag{1.271}$$

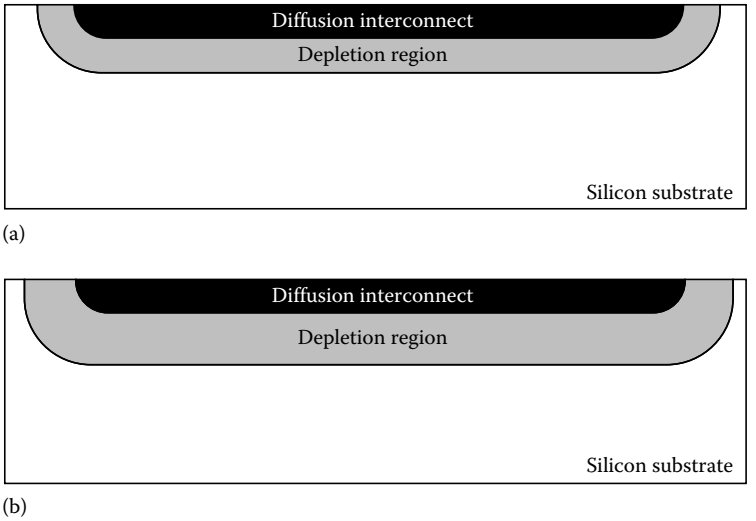


FIGURE 1.107 Diffusion interconnect has a voltage-dependent capacitance produced by the depletion region between the interconnect and the substrate. At low voltage difference between the interconnect and substrate (a), the capacitance is large. However, the capacitance decreases for larger voltage differences (b).

The typical equation for depletion capacitance is given in Equation 1.271. In this equation V_S is the voltage from the interconnect to the substrate, ϕ_B is the built-in potential of the semiconductor junction, M is the grading coefficient of the junction, while C_0 is the zero-bias capacitance of the junction. Since the capacitance is less than C_0 for reverse bias and the junction would not insulate for forward bias, we can assume that the capacitance is always less than C_0 and use C_0 as a conservative estimate of C. Because of the uncertainty in the exact structure of most semiconductor junctions φ_B and M are usually fit to measured capacitance versus voltage (CV) data.

Another common nonlinear parasitic occurs when metal interconnect placed over a conducting semiconductor substrate creates inversions at the semiconductor surface. This inversion layer increases the substrate capacitance of the interconnect and is voltage-dependent. To prevent this most silicon-IC manufacturers place an inversion-preventing implant on the surface of the substrate. The depletion between the substrate and n-type or p-type wells diffused into the substrate also creates a voltage-dependent capacitance. Thus use of the well as a high resistivity interconnect for making high value resistors will require consideration of a nonlinear capacitance to the substrate.

1.5.2 Pad and Packaging Parasitics

All signals and supply voltages that exit an IC must travel across the packaging interconnections. Just like the on-chip interconnect, the packaging interconnect has parasitic resistance, capacitance, and inductance. However, some of the packaging materials are significantly different in properties and dimension to those used in the IC, thus there are major differences in the importance of the various types of parasitics. Figure 1.108 is a typical packaged IC. The chief components of the packaging are the pads on the chip, the wire or bump bond used to connect the pad to the package, and then the package interconnect.

The pads used to attach wire bonds or bump bonds to ICs are often the largest features on an IC. The typical pad is 100 μm on a side and has a capacitance of 100 fF. In addition, protection diodes are often used on pads that will add a small nonlinear component to the pad capacitance.

The wire bonds that attach the pads to the package are typically very low resistivity and have negligible capacitance. Their major contribution to package parasitics is inductance. Typically, the

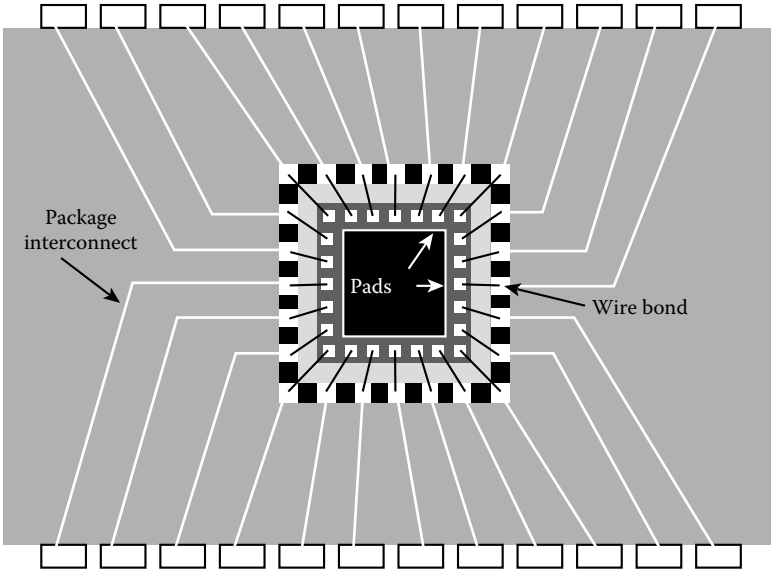


FIGURE 1.108 A packaged IC. The main sites of parasitics are the pad, bond, and package interconnect.

package interconnect inductance is greater than the wire bond inductance; however, when wire bonds are used to connect two ICs directly together, then the wire bond inductance is significant.

Often, the dominant component of package parasitics comes from the packaging interconnect itself. Depending on the package, there is inductance, capacitance to ground, and parallel line capacitance produced by this interconnect. Carefully made high-frequency packages do not exhibit much parallel line capacitance (at the expense of much capacitance to ground due to shielding), but in low-frequency packages with many connections this can become a problem.

Typical inductance and capacitance values for a high-speed package capable of output bandwidths around 5 GHz are incorporated into a circuit model for the package parasitics in Figure 1.109. When simulated with a variety of circuit source resistances (*RS*) this circuit reaches maximum bandwidth without peaking when the output resistance is 4 Ω. At lower output resistance, Figures 1.110 and 1.111 show that considerable peaking in the output frequency response occurs.

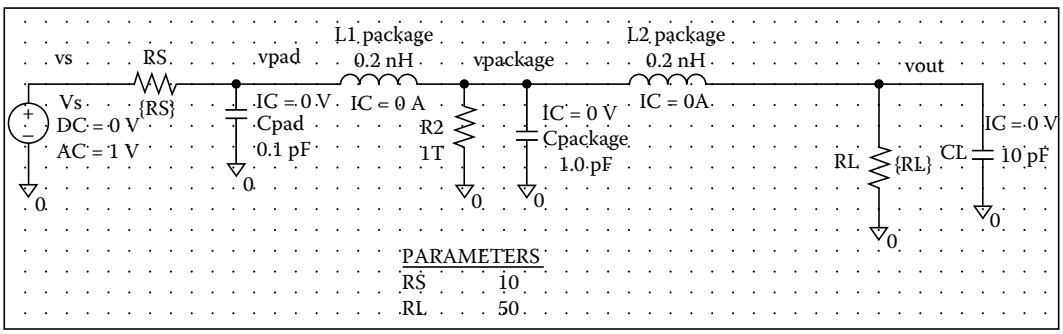


FIGURE 1.109 The circuit model of a high-frequency package output and associated parasitics. Cpad is the pad capacitance. L1package, Cpackage, R2, and L2package model the package interconnect. *RS* is the source resistance of the circuit. *RL* and *CL* are the external load.

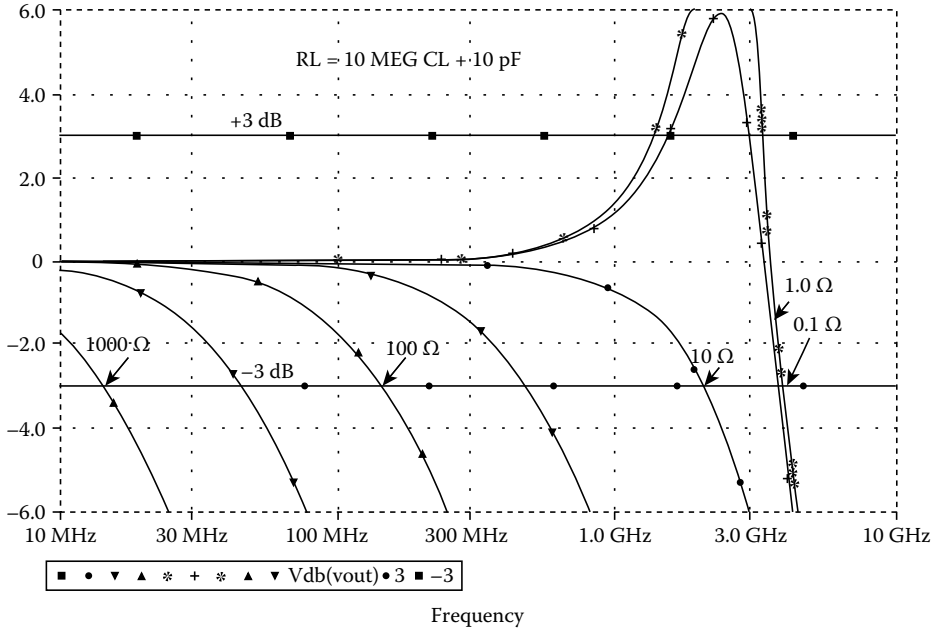


FIGURE 1.110 The PSPICE ac simulation of circuit in Figure 1.109 when the load resistance *RL* is 10 MΩ. This shows how the package inductance causes peaking for sufficiently low output resistance. In this case, peaking occurs for *RS* below 4 Ω and at about 2 GHz.

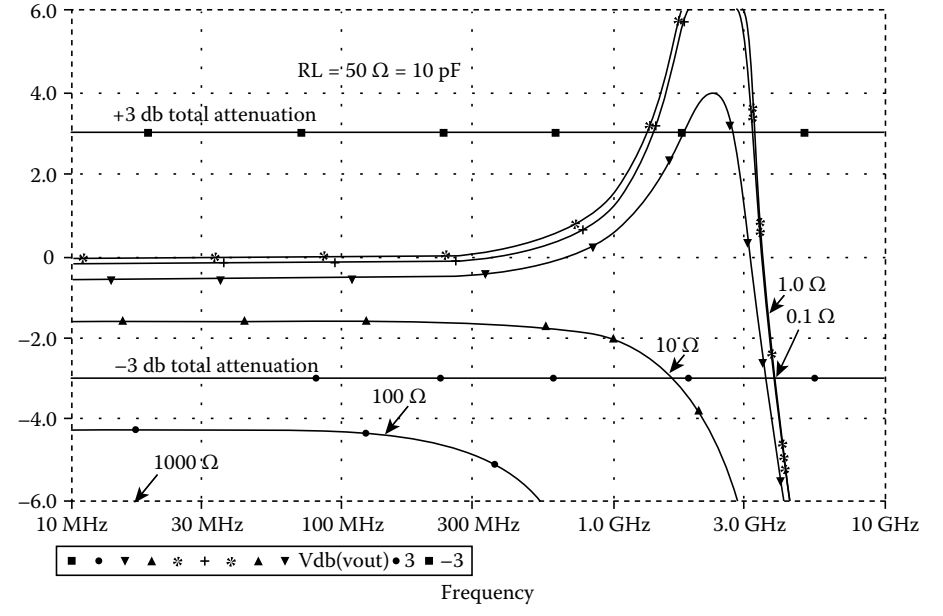


FIGURE 1.111 The PSPICE ac simulation of circuit in Figure 1.109 when the load is 50 Ω. The package inductance still causes peaking for *RS* below 4 Ω.

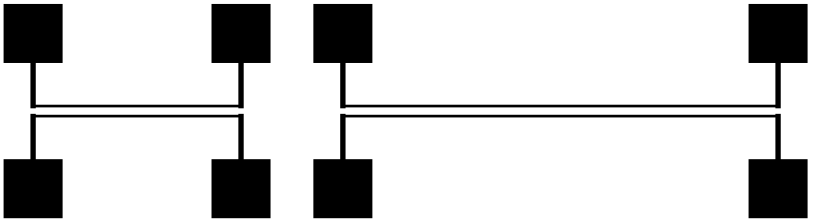


FIGURE 1.112 Test structures for measuring parallel line capacitance.

1.5.3 Parasitic Measurement

The major concern when measuring parasitics is to extract the individual parasitic values independently from measured data. This is normally achieved by exaggerating the effect that causes each individual parasitic in a special test structure, and then reproducing the structure with two or more different dimensions that will affect only the parasitic of interest. In this fashion, the effects of the other parasitics are minimized and can be subtracted from the desired parasitic in each measurement.

$$C_P = \frac{C_1 - C_2}{L_1 - L_2} \tag{1.272}$$

For example, to measure parallel line capacitance, the test structures in Figure 1.112 would be fabricated. These structures vary only in the length of the parallel lines. This means that if other parasitic capacitance ends up between the two signal lines used to measure the parasitic, then it will be a constant capacitance that can be subtracted from both measurements. The parallel line capacitance will vary in proportion to the variation of length between the two test structures. Thus the parallel line capacitance per unit length can be found from Equation 1.272. In this equation C_P is the parallel line capacitance per unit length, C_1 and C_2 are the capacitances measured from each test structure, and L_1 and L_2 are the length of the two parallel interconnect segments.

References

1. D. L. Carter and D. F. Guise, Effects of interconnections on submicron chip performance, *VLSI Design*, 4, 63–68, 1984.
2. H. B. Lunden, Detailed extraction of distributed networks and parasitics in IC designs, in *Proc. Euro. Conf. Circuit Theory, Design*, 1989, pp. 84–88.
3. R. A. Sainati and T. J. Moravec, Estimating high-speed circuit interconnect performance, *IEEE Trans. Circuits Syst*., 36, 533–541, April 1989.
4. D. S. Gao, A. T. Yang, and S. M. Kang, Modeling and simulation of interconnection delays and crosstalks in high-speed integrated circuits, *IEEE Trans, Circuits Syst*., 37, 1–8, Jan. 1990.
5. M. Horowitz and R. W. Dutton, Resistance extraction from mask layout data, *IEEE Trans. Computer-Aided Design Integrat. Circuits Syst*., 7, 1029–1037, Oct. 1988.
6. R. J. Antinone and G. W. Brown, The modeling of resistive interconnections for integrated circuits, *IEEE J. Solid-State Circuits*, SC-18, 200–203, April 1983.
7. A. E. Ruehli and P. A. Brennan, Capacitance models for integrated circuit, metallization wires, *IEEE J. Solid-State Circuits*, SC-10, 530–536, Dec. 1975.
8. S. Mori, I. Suwa, and J. Wilmore, Hierarchical capacitance extraction in an IC artwork verification system, in *Proc. IEEE Int. Conf. Computer-Aided Design*, 1984, pp. 266–268.

2

Analog Circuit Cells

Kenneth V. Noren
University of Idaho

John Choma, Jr.
University of Southern California

J. Trujillo
University of Southern California

David G. Haigh
University College of London

Bill Redman-White
University of Southampton

Rahim Akbari-Dilmaghani
University College of London

Mohammed Ismail
Ohio State University

Shu-Chuan Huang
Ohio State University

Chung-Chih Hung
Tatung Institute of Technology

Trond Saether
Nordic VLSI A/S

2.1 Bipolar Biasing Circuits **2**-1
Common Bipolar Junction Transistor (BJT) Biasing Circuits
References **2**-10
2.2 Canonic Cells of Linear Bipolar Technology **2**-10
Introduction • Small-Signal Model • Single-Input–Single-Output Canonic Cells • Differential Amplifier
References **2**-54
2.3 MOSFET Biasing Circuits **2**-55
Introduction • Device Types and Models for Biasing • Voltage and Current Reference and Bias Circuits • Voltage and Current References Based on Less Usual Devices • Voltage References Based on N- and P-Doped Polysilicon Gate Threshold • Biasing of Simple Amplifiers and Other Circuits • Biasing of Circuits with Low Power Supply Voltage • Dynamic Biasing • Conclusions
References **2**-76
2.4 Canonical Cells of MOSFET Technology **2**-77
Matched Device Pairs • Unmatched Device Pairs • Composite Transistors • Super MOS Transistors • Basic Voltage Gain Cells • Conclusion
References **2**-98

2.1 Bipolar Biasing Circuits

Kenneth V. Noren

Establishing bias currents and voltages for building blocks comprising an overall design is fundamental to the design of bipolar integrated circuits. These building blocks include single-stage and differential amplifiers, output stages, etc. Biasing often has a direct relationship to electrical characteristics, such as gain, signal-swing, slew-rate, etc., of the individual building blocks and hence to the overall design. Biasing circuits include current sources, voltage references, and level-shifters. Most often, it is desirable that the integrated circuit design be robust and independent of a variety of external factors that can affect circuit performance. These factors include variations in process parameters, supply voltage, and temperature. Efforts to improve the performance of current sources and voltage references have led to many refinements and developments that have started from simple beginnings. This section presents some of

the fundamental current sources and voltage references used for biasing in bipolar integrated circuit technologies and refinements of these circuits that have evolved over time.

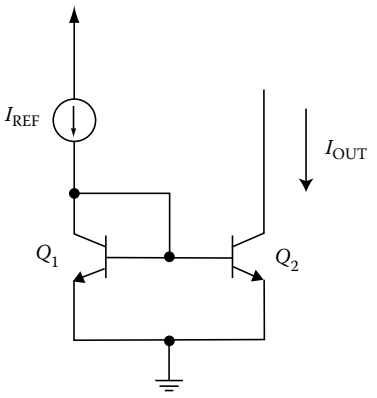


FIGURE 2.1 Simple current mirror.

2.1.1 Common Bipolar Junction Transistor (BJT) Biasing Circuits

2.1.1.1 Current Mirrors and Sources

The current mirror is a circuit that reproduces a reference current at one or more locations in larger circuit. A simple current mirror is depicted in Figure 2.1. Since $V_{BE1} = V_{BE2}$, $I_{OUT} \approx I_{REF}$, and the reference current I_{REF} is effectively mirrored to another location. In order to evaluate current mirrors and compare the properties of the many types of current mirrors to one another, we first define metrics for current mirrors and characteristics for an ideal current mirror. An ideal current mirror produces and output current that

1. Reproduces a reference current, exactly
2. Does not vary with loading (the output resistance [R_o] is infinite)
3. Is insensitive to process variations
4. Is insensitive to power supply variations
5. Is insensitive to temperature

2.1.1.1.1 Simple Current Mirror

The relationship between the output current and the reference current or the simple current mirror for matched transistors is

$$\frac{I_{OUT}}{I_{REF}} = \frac{1}{1 + (2/\beta)} \tag{2.1}$$

This equation does not include the effects of the early voltage, but does include the effects of nonzero base current often referred to as errors due to finite β. The error due to finite β results because I_{REF} must supply base current to Q_1 and Q_2. The key problem with this dependency of I_{OUT} on β is that β may vary from due to process variations, resulting in an I_{OUT} that varies due to process variations. Were it not for this dependency, I_{REF} could be adjusted to compensate for this and set I_{OUT} to a desired value. For $\beta \gg 2$, the fractional error is $-2/\beta$ percent.

A second error in I_{OUT} occurs due to finite output resistance. Performing a small-signal analysis for the circuit for Figure 2.1, it can be shown that the expression for R_o for the simple current source is equal to r_{o2}.

If the mirror in Figure 2.1 is configured with N output transistors and thus has N output currents, the transfer function becomes

$$\frac{I_{OUT}}{I_{REF}} = \frac{1}{1 + ((N+1)/\beta)} \tag{2.2}$$

2.1.1.1.2 Simple Current Mirror with Beta Helper

The β sensitivity of the simple current mirror can be improved by adding a third transistor to supply base current to Q_1 and Q_2 shown in Figure 2.2. Here, the base current of Q_1 and Q_2 is supplied by the emitter of Q_3 which draws the necessary current from I_{REF}, but reduced by a factor of $\beta + 1$. An analysis shows

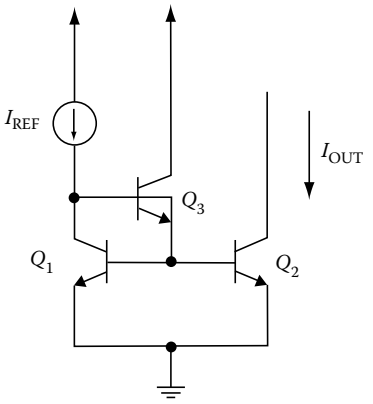


FIGURE 2.2 Simple current mirror with beta helper

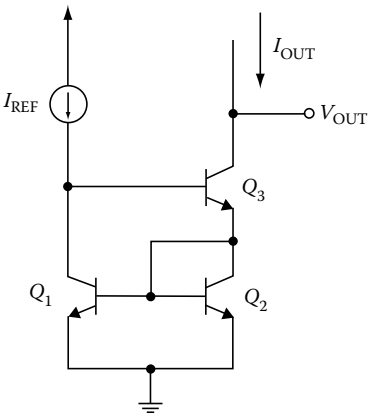


FIGURE 2.3 Wilson current mirror.

$$\frac{I_{OUT}}{I_{REF}} = \frac{1}{1 + (2/(\beta^2 + \beta))} \approx \frac{1}{1 + (2/\beta^2)} \tag{2.3}$$

This supports the argument that sensitivity to base current drain is reduced. It should be noted that the equation also depends on the betas of the transistors being matched as well. For $\beta \gg 2$, the fractional error is $-2/\beta^2$ percent. Thus, the fractional error is reduced by a factor of β. For this circuit $R_o = r_{o2}$, so there is no improvement in R_o.

2.1.1.1.3 Wilson Current Mirror

A current source that shows an improvement in R_o and has reduced β sensitivity is the Wilson current mirror shown in Figure 2.3 [1].

With this circuit, as with the simple current mirror with beta helper, the base current of Q_1 and Q_2 is supplied by emitter current of a third transistor, Q_3. A more rigorous analysis will show, neglecting the effects of the early voltage,

$$\frac{I_{OUT}}{I_{REF}} = \frac{1}{1 + (2/(\beta^2 + \beta))} \approx \frac{1}{1 + (2/\beta^2)} \tag{2.4}$$

For $\beta \gg 2$, the fractional error is $-2/\beta^2$. Thus, the fractional error is reduced by a factor of β.

The improvement in R_o is due to negative feedback present in the circuit due to the placement of Q_3. To see this, first consider the case for matched transistors and the effects of output voltage for all of the transistors, that an increase in V_{OUT} gives rise to an increase in output current. This, in turn, causes an increase in I_{C2}. Since Q_1 and Q_2 themselves form a simple current mirror, I_{C1} also increases which forces a decrease in I_{B3}, since I_{REF} is constant. This in turn reduces in I_{OUT}. A small-signal analysis to determine R_o shows

$$R_o \approx \frac{\beta}{2} r_o \tag{2.5}$$

R_o has been increased by a factor of $\beta/2$.

The Wilson current mirror can also be extended to N multiple outputs by placing additional transistors branches in parallel with Q_2 and Q_3. For this case, it can be shown that

$$\frac{I_{OUT}}{I_{REF}} = \frac{1}{N} \frac{1}{1 + (2/\beta^2)} \tag{2.6}$$

where N is the total number of output branches.

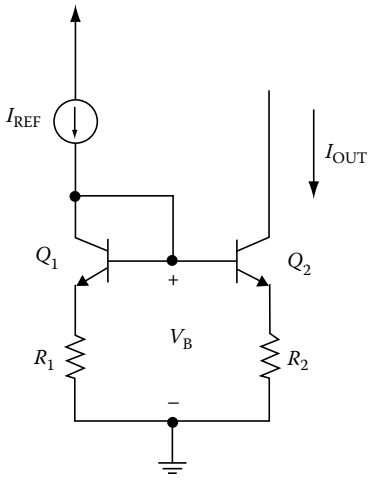


FIGURE 2.4 Simple current mirror with emitter degeneration.

2.1.1.1.4 Simple Current Mirror with Emitter Degeneration

Parameters and component values for fabricated devices usually exhibit deviation from some "nominal" value during the fabrication process. These variations may occur for across the die, from die-to-die, from wafer-to-wafer, or from lot-to-lot. An objective in designing circuits that are process insensitive is to minimize the effects of process variation. A good example of this is the simple current mirror with emitter degeneration shown in Figure 2.4. If I_{REF} and R_1 are such that the voltage drop across V_{BE1} is small in comparison, then the dominant voltage at V_B is approximately $I_{REF}R_1$. Likewise, if R_2 and I_{OUT} are such that the voltage drop across V_{BE2} can be neglected, then we have $I_{REF}R_1 \approx I_{OUT}R_2$ and

$$I_{OUT} = \frac{R_2}{R_1} I_{REF} \tag{2.7}$$

For the simple current mirror, neglecting β and considering the possibly of mismatched emitter areas, we have

$$I_{OUT} = \frac{A_2}{A_1} I_{REF} \tag{2.8}$$

Since resistors can be matched to ±0.1% and NPN matching for transistors can be as poor as ±1%, the current mirror with emitter degeneration is less susceptible to processing errors. The process insensitivity is made possible by having the relationship between I_{OUT} and I_{REF} dependent on resistor ratios.

2.1.1.1.5 Widlar Current Mirror

In bipolar integrated circuit design, it is sometimes desirable to create low currents levels, on the order of microamps, for example [2]. If either of the simple current mirrors or the Wilson current mirror is used, this has to be accomplished by creating a very small reference current and may require large values of resistors that may consume large amounts of area. The current mirror depicted in Figure 2.5 is capable of producing a small output current, from a nominal reference current and a reasonably sized resistor.

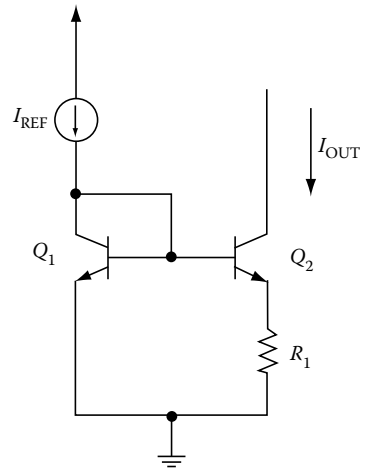


FIGURE 2.5 Widlar current mirror.

To analyze the circuit, recognize that I_{REF} is determines V_{BE1}. In the simple current mirror, all of V_{BE1} appears across V_{BE2}. In the Widlar current mirror, V_{BE1} is divided between the base–emitter junction of Q_2 and R_1, resulting in a smaller voltage for V_{BE2} than V_{BE1} and thus a smaller I_{OUT}. This suggests that with proper selection of R_1, the potential for generating very small currents exists. It can be shown, neglecting the effects of β, that

$$I_{OUT}R_1 = V_T \ln\left(\frac{I_{REF}}{I_{OUT}}\right) \tag{2.9}$$

where V_T is the thermal voltage.

For example, to create an I_{OUT} of 5 μA, from a reference of 1 mA, with $V_T = 26$ mV, we find $R_1 = 27$ kΩ. If the same 5 μA were desired from the simple current mirror, and we assume that the emitter areas of Q_1 and Q_2 are equal, we would need to generate a reference current of 5 μA. To do this, I_{REF} is replaced by a resistor, R_{REF}, tied to the positive supply voltage, V_{CC}, for example. Suppose for this example that V_{CC} is 5 V. Then, the voltage drop across R_{REF} is $V_{CC} - V_{BE}$. Taking $V_{BE} = 0.7$ V gives a value of $R_{REF} = (5 - 0.7)/5\mu = 860$ kΩ. This, of course, takes up much more chip real estate than for $R_1 = 27$ kΩ and is undesirable.

If the effects of β are included it is necessary to supply two base currents, though I_{B2} is less than I_{B1}. Equation 2.2 gives an upper bound for the error due to beta for the Widlar current mirror (the upper bound being the case where $I_{OUT} = I_{REF}$) and lower bound being $1/(1 + 1/\beta)$, supplying base current to only a single transistor. Thus, the errors due to base current drain are on the same order as that of the simple current mirror.

For the output resistance, R_1 provides negative feedback and thus increases the output resistance compared with that of the simple current mirror. It is identical to the increase in output resistance that results from emitter degeneration in a common-emitter amplifier. It can be verified that

$$R_o = (1 + g_m r_\pi R_1) r_o \tag{2.10}$$

2.1.1.1.6 Low-Bias Current Mirror

An alternative to the Widlar current source that also provides a low output current is the current source shown in Figure 2.6 [3,4]. Again, V_{BE1} is determined by I_{REF}. Applying Kirchhoff's voltage law,

$$V_{BE2} = V_{BE1} - I_{REF} R_1 \tag{2.11}$$

As with the Widlar current source voltage, the voltage V_{BE1} is divided between a resistor and the base emitter junction of Q_2, though in a less obvious manner. As a result, V_{BE2} must be smaller than V_{BE1} and effectively a fraction of I_{REF} is mirrored. A more exact equation that expresses the relationship between the output current and reference current can be derived from Equation 2.8, and this is

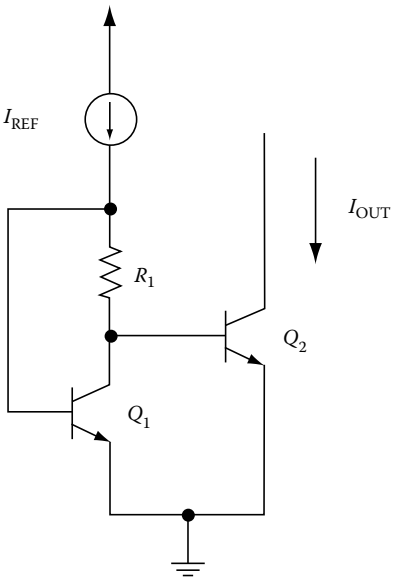


FIGURE 2.6 Current mirror for generating low-bias currents.

$$I_{OUT} = \frac{I_{REF}}{\exp(I_{REF} R_1 / V_T)} \tag{2.12}$$

or

$$R_1 = \frac{V_T \ln(I_{REF}/I_{OUT})}{I_{REF}} \tag{2.13}$$

In fact, this current source is capable of supplying even lower currents than the Widlar current source for a given I_{REF} and a lower bound for R_1. Consider the same example as was given for the Widlar current source. With $I_{REF} = 1$ mA, $I_{OUT} = 5$ μA, and $V_T = 26$ mV, we find that $R_1 = 137.75$ Ω. This is a substantial decrease in resistance from the Widlar current mirror example.

The price paid for this improvement is a reduced output resistance when compared to the Widlar current

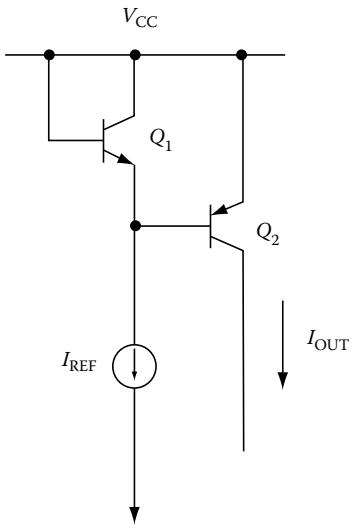


FIGURE 2.7 Current mirror for complementary bipolar design.

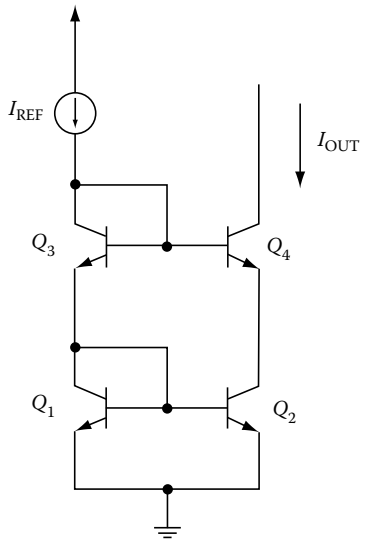


FIGURE 2.8 Cascode current mirror.

source. Negative feedback is not present in this topology and there are no improvements in output resistance when compared to the simple current mirror. For this current mirror, R_o is simply r_{o2}.

The basic current mirrors can be extended to complementary bipolar technology (CBT) as well. Figure 2.7 shows an example of a current mirror that can be found in a CBT. A problem that arises in the current mirrors used in bipolar technology results in the fact that PNP and NPN transistors have a different Gummel number [5]. This results in different base–emitter voltages (magnitudes) for equal collector currents. The effects of this can be deduced from Figure 2.7, where now ΔV_{BE} errors must be considered. Thus, care must be taken when biasing complementary bipolar designs. This issue of balancing and matching is a fundamental problem in this technology [5].

2.1.1.1.7 Cascode Current Mirror

The cascode current mirror is depicted in Figure 2.8. It derives its main advantage in an increased output resistance due to emitter degeneration as does the Widlar current source. In this case, r_{o2} replaces R_1 in the Widlar current source to provide the emitter degeneration. Though in theory the values of r_{o2} and R_1 may be on the same order, a large value for r_{o2} can be achieved using a transistor that takes up much less area than that of a resistor of the same value.

The complete expression for R_o is complicated. However, for the case where $I_{OUT} \approx I_{REF}$, the transistors are matched and $g_m r_o \gg 1$, $\beta \gg 2$, and $r_o \gg r_\pi$ (for any combination of transistors) the expression for the output resistance reduces to

$$R_o = \frac{\beta}{2} r_{o2} \tag{2.14}$$

2.1.1.1.8 V_{BE} Referenced Current Mirror

For many current mirrors, I_{REF} is determined by a resistor tied to the positive power supply. In the Widlar current mirror example, $I_{REF} = (V_{CC} - V_{BE})/R_{REF}$. The reference current is directly proportional to the supply voltage. In many situations, this is undesirable. One alternative is to replace V_{CC} by one of

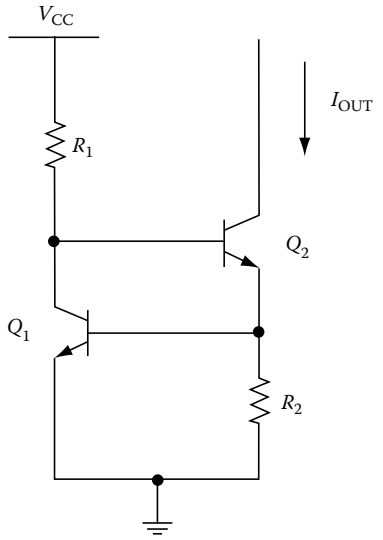


FIGURE 2.9 Current source that has reduced supply voltage dependency.

the many available circuits which provide a voltage reference that is independent of supply voltage. Another alternative is the V_{BE} referenced current mirror (Figure 2.9).

The basic principle of this current mirror is to establish a base–emitter voltage and to convert this voltage to a current using a resistor. Referring to Figure 2.9, the V_{BE} (V_{BE1}) drop is first established with V_{CC} and R_1. The voltage V_{BE1} and R_2 determines I_{OUT}. Neglecting finite betas for the transistors, this current is approximately equal to V_{BE1}/R_2. Since V_{BE1} is fairly constant, I_{OUT} is fairly constant. A complete derivation yields

$$I_{OUT} = \frac{V_T}{R_2} \ln\left(\frac{V_{CC}}{I_{S1}R_1}\right) \tag{2.15}$$

where I_{S1} is the saturation current of Q_1. Equation 2.13 shows the output current has a logarithmic variation with respect to V_{CC}, an improvement over the linear relationship found in other current mirrors.

2.1.1.1.9 Self-Biased V_{BE} Referenced Current Source

A self-biased V_{BE} referenced current source is depicted in Figure 2.10. Q_1, Q_2, Q_3, Q_4, and R_2 form the core of the current source. Q_1 and Q_2 form a V_{BE} referenced current mirror and the pair Q_3 and Q_4 form a simple current mirror. If current exists, then $I_{C1} = I_{C2}$, due to Q_3 and Q_4, and one valid solution to is $I_{C2} \approx V_{BE}/R_2$, independent of V_{CC}. However, a second valid solution is $I_{C2} = 0$ A (or practically, a value for I_{C2} on the order of leakage currents). For this reason, a start-up circuit is added to the circuit. D_{1-5}, R_{B1}, and R_{B2} form the start-up circuitry. If $I_{C1} = I_{C2} = 0$, the voltage at the cathode of D_1 is 0 V and D_1 turns on, injecting current into the core of the current source. Positive feedback in the circuit forces the current toward the condition where $I_{C2} \approx V_{BE}/R_2$. At some point, the voltage at the cathode of D_1 raises to a level that shuts D_1 off, thereby "disconnecting" the start-up circuitry from the current source core.

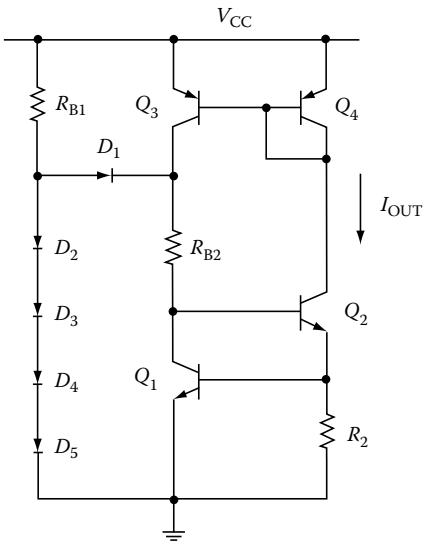


FIGURE 2.10 Self-biased V_{BE} referenced current source.

Once I_{C1} and I_{C2} have been established as a reference current for a larger circuit, the current can be mirrored to other parts of the circuit by placing transistors in parallel with Q_1 and Q_4, as shown with Q_5 and Q_6.

2.1.1.1.10 Self-Biased V_T Referenced Current Source

A second type of self-biased current source called self-biased V_T referenced current source is shown in Figure 2.11. In this circuit, Q_3 and Q_4 are current mirrors and force the condition that if current exists, then $I_{C1} = I_{C2}$. In general, the area of Q_2 (A_2) is set to be "N" times the area of Q_1 (A_1) and thus $I_{S2} = NI_{S1}$, where n is some integer. In practice, this is achieved by placing n transistors in parallel. This causes a difference in base–emitter voltages, $\Delta V_{BE} = V_{BE1} - V_{BE2} = nV_T \ln(N)$, and this difference is dropped across R_1. Thus,

$$I_{C1} = I_{C2} = \frac{nV_T \ln(N)}{R_1} \tag{2.16}$$

I_{C1} and I_{C2} can in turn be mirrored by placing transistors in parallel with Q_1 and Q_4 to form current mirrors.

2.1.1.2 Voltage References

Also fundamental to biasing in BJT circuits is the voltage reference. As with the current sources, we may define an ideal voltage source in order to have an adequate way of evaluating voltage references. An ideal voltage reference produces a voltage that

1. Does not vary with loading (zero output resistance)
2. Is insensitive to process variations
3. Is insensitive to power supply variations
4. Is insensitive to temperature

The simplest voltage reference is the zener diode reference, but it is well known that a zener diode exhibits temperature dependence and has a fairly high output resistance. Some of the modifications to a simple zener diode include placing a diode series with a zener, strings of diodes, and the common-collector stage. Figure 2.12 gives an example. Here it is assumed, but not always true, that V_{DZ} has a positive coefficient and V_{D1} has a negative temperature coefficient and provides some cancellation of the coefficients. Thus, this can produce a reference voltage that is insensitive to temperature. There are literally hundreds of deviations based on this fundamental principle.

A circuit that produces a voltage that is an arbitrary multiple of V_{BE} is the V_{BE} multiplier circuit shown in Figure 2.13. The circuit works on the principle that a current equal to V_{BE}/R_2 is generated and, neglecting base current, flows through R_1. Thus, the voltage across R_1 is $(V_{BE}/R_2)R_1$ and the total voltage can be written as

$$V_{REF} = \left(1 + \frac{R_2}{R_1}\right) V_{BE} \tag{2.17}$$

Many applications for biasing circuits demand that their performance remain constant through a wide range of temperatures. Thus, many circuits for temperature insensitive biasing have emerged. Ideally, a temperature insensitive output, voltage or current, would depend on a temperature insensitive element. Since all semiconductor components exhibit variation with temperature, most of the schemes for temperature independence involve some form of cancellation technique or compensation [6]. Instead of eliminating all of the sensitivity, the design techniques strive to minimize the errors. For example, a common solution is to place devices with positive temperature coefficients in series with devices with negative temperature coefficients, scaling some of these coefficients if necessary, to provide nearly zero sensitivity to temperature for an output is taken across the devices.

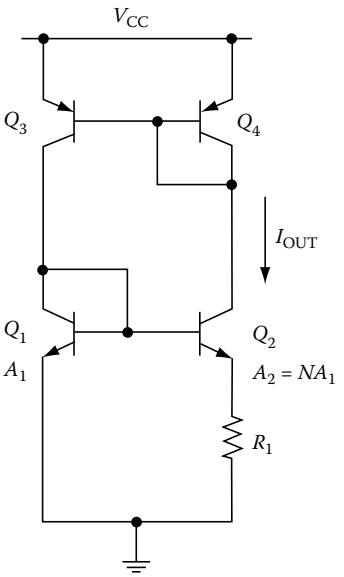


FIGURE 2.11 Self-biased VT referenced current source.

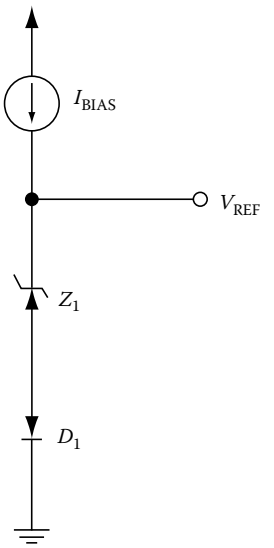


FIGURE 2.12 Zener-biased voltage reference.

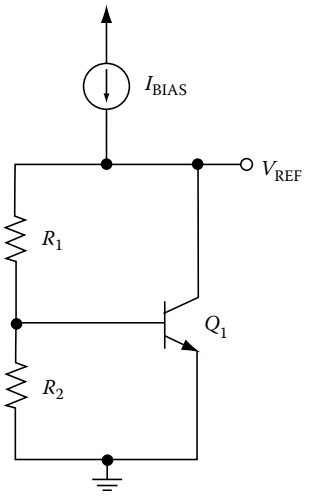


FIGURE 2.13 V_{BE} multiplier circuit.

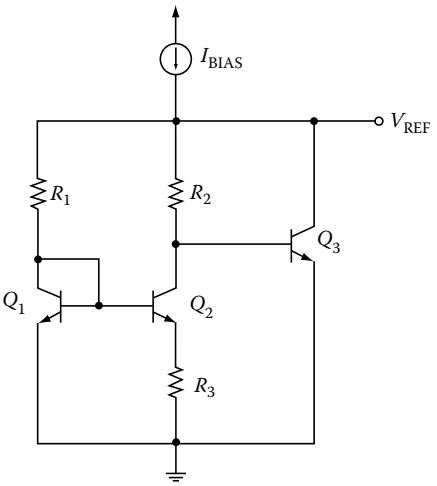


FIGURE 2.14 Simple bandgap reference.

A simple band-gap reference is depicted in Figure 2.14. Band-gap circuits also operate on a principle of cancellation of temperature coefficients. Generally, a voltage is developed which is a scaled value of V_T. This scaled value has a well-defined temperature coefficient which is the scaling constant times the temperature coefficient of V_T, which is positive. This voltage is added to a base–emitter voltage, which has a negative temperature coefficient. The scaling factor is chosen so that the sum of the temperature coefficients is zero. The output is then taken across the two voltages to produce a voltage with a temperature coefficient of approximately zero. Generally, the output voltage has the form

$$V_{REF} = V_{BE} + KV_T \tag{2.18}$$

If the temperature coefficient of V_{BE} is taken to be -2 mV/°C and the temperature coefficient of V_T is taken to be $k/q \approx +0.085$ mV/°C, this results in a value for K of about 23.52. This gives a value for V_{OUT} of about 0.7 V + 23.52(25.9 mV) = 1.3 V, close to the band-gap voltage of silicon, and gives rise to the nomenclature of band-gap references.

In the circuit in Figure 2.14, we assume that Q_1 and Q_2 operate at different current densities. This is done either by operating Q_1 and Q_2 at different collector current levels, with matched emitter areas, or by operating them at the same collector currents with emitter areas being mismatched. A voltage ΔV_{BE} is developed across R_3 and then current through R_3 and R_2, negecting the effects of β, is $\Delta V_{BE}/R_3$. This gives $V_{REF} = \Delta V_{BE} R_2/R_3 + V_{BE3}$. Since $\Delta V_{BE} = V_T \ln(J_1/J_2)$, this gives the constant K as being $R_2/R_3 \ln(J_1/J_2)$.

An improved band-gap reference is depicted in Figure 2.15. This reference forms a basic building block for several commercial voltage references. In this circuit, I_{C1} and I_{C2} are forced to be equal by the high-gain

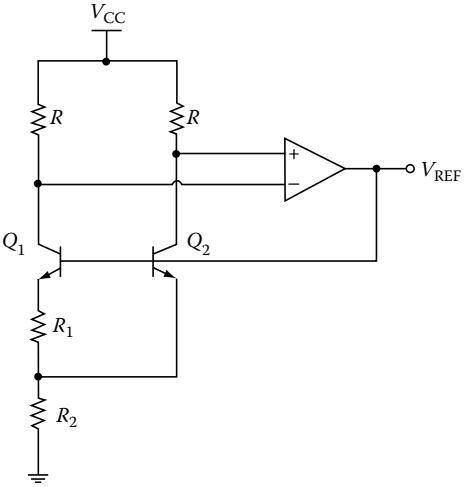


FIGURE 2.15 Improved bandgap reference.

amplifier operating with negative feedback. The current densities of Q_1 and Q_2 are made unequal by sizing the emitter areas of Q_1 and Q_2 differently. In this case, $A_1 = NA_2$. This means $\Delta V_{BE} = nV_T\ln(I_{S1}/I_{S2}) = V_T\ln(N)$. The voltage drop across R_2 is $\Delta V_{BE}R_2/R_1$. Finally, $V_{REF} = V_{BE} + V_T\ln(n)R_2/R_1$. Thus, in this case, $K = R_2/R_1\ln(n)$.

There is a wealth of information available in the literature on current sources and voltage references. Further depth into bias circuits behavior is provided in the accompanying references for this section.

References

1. G. R. Wilson, A monolithic junction FET-NPN operational amplifier. *IEEE J. Solid-State Circuits*, SC-3, 341–348, Dec. 1968.
2. R. J. Widlar, Some circuit design techniques for linear integrated circuits. *IEEE Trans. Circuit Theory*, CT-12, 586–590, Dec. 1965.
3. C. Kwok, Low-voltage peaking complementary current generator. *IEEE J. Solid-State Circuits*, SC-20, 816–818, Jun. 1985.
4. P. R. Gray, P. J. Hurst, S. H. Lewis, and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 4th ed., New York: Wiley, 2001, pp. 253–334.
5. C. Toumazou, F. J. Lidgey, and D. G. Haigh (ed.), *Analogue IC Design: The Current-Mode Approach*, London: Peter Peregrinus, 1990, Chapters 6 and 16.
6. A. Grebene, *Bipolar and MOS Analog Integrated Circuit Design*, New York: Wiley, 1984, Chapter 4.

2.2 Canonic Cells of Linear Bipolar Technology

John Choma, Jr. and J. Trujillo

2.2.1 Introduction

The circuit configurations of linear signal processors realized in bipolar technology are as diverse as the system operating requirements that these circuits are designed to satisfy. Despite topological diversity, most practical open-loop linear bipolar circuits are derived from interconnections of surprisingly few basic subcircuits. These subcircuits include the diode-connected bipolar junction transistor (BJT), the common-emitter amplifier, the common-base amplifier, the common-emitter–common-base cascode, the emitter follower, the Darlington connection, and the balanced differential pair. Because these open-loop subcircuits underpin linear bipolar circuit technology, they are rightfully termed the "canonic cells" of linear bipolar circuit design.

By examining the low-frequency performance characteristics of the canonic cells of linear bipolar technology, this section achieves two objectives. First, the forward gain, the driving point input resistance, and the driving point output resistance are cataloged for each canonic circuit. This information produces Thévenin and Norton I/O port equivalent circuits that expedite the analysis and design of multistage electronics. Second, the forthcoming work establishes a basis for prudent circuit design in that all analytical results are studied by highlighting the attributes and uncovering the limitations of each cell. The understanding that resultantly accrues paves the way toward systematic design procedures that yield optimal circuit architectures capable of circumventing observed subcircuit shortcomings.

2.2.2 Small-Signal Model

The fundamental tool exploited in the analyses that follow is the low-frequency small-signal equivalent circuit of a BJT shown in Figure 2.16a. This equivalent circuit, which applies to NPN and PNP discrete component and monolithic transistors, is derived from the low-frequency, large-signal NPN BJT model

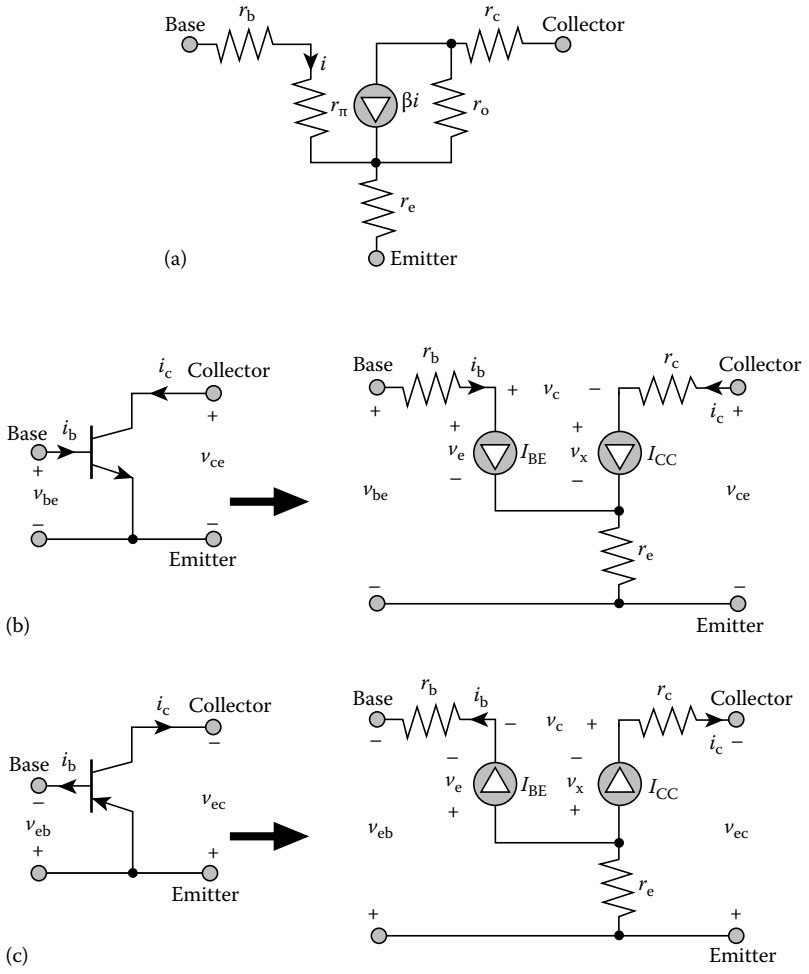


FIGURE 2.16 (a) Low-frequency, small-signal model of a BJT. (b) Low-frequency, large-signal model of an NPN BJT. (c) Low-frequency, large-signal model of a PNP BJT.

offered in Figure 2.16b [1,2]. As is depicted in Figure 2.16c, the PNP large-signal transistor model is topologically identical to its NPN counterpart. The only difference between the two models is a reversal in the direction of all controlled current sources and branch currents and a reversal in polarity of all assigned branch and port voltages.

The large-signal models in Figure 2.16b and c are simplified to reflect transistor biasing that assures nominally linear device operation for all values of applied input signal voltages. A necessary condition for linear operation is that the internal emitter–base junction voltage v_e be at least as large as the threshold voltage, say v_γ, of the junction for all time, that is,

$$v_e(t) \geq v_\gamma \quad \text{for all time } t \tag{2.19}$$

For silicon transistors, v_γ is typically in the neighborhood of 700–750 mV. A second condition underlying transistor operation in its linear regime is that the internal base–collector junction voltage v_c is never positive, that is,

$$v_c(t) \leq 0 \quad \text{for all time } t \tag{2.20}$$

In the models of Figure 2.16b and c, r_b represents the "effective base resistance" of a BJT, r_c is its net "internal collector resistance," and r_e is the net "internal emitter resistance." All three resistances, and particularly r_b, decrease monotonically with increasing quiescent base and collector currents, I_{BQ} and I_{CQ}, respectively [3,4]. The collector resistance also decreases with increase in the intrinsic collector–emitter voltage v_x. Large base, collector, and emitter resistances conduce reduced circuit gain, diminished gain-bandwidth product, and increased electrical noise. In view of these observations and in the interest of formulating a mathematically tractable analysis that produces conservative estimates of bipolar circuit performance, these resistances are usually interpreted as constants equal to their respective low-current, low-voltage values.

In a monolithic fabrication process, unacceptably large internal device resistances can be reduced by exploiting the fact that r_b, r_c, and r_e are inversely proportional to the emitter–base junction injection area. This area is a designable parameter chosen to ensure that the transistor in question conducts the proper density of collector current. Unfortunately, the engineering price potentially paid for a reduction of device resistances through increase in junction area is circuit response speed, since the capacitances associated with transistor junctions are directly proportional to device injection area.

The current I_{BE} in Figure 2.16b and c is given approximately by

$$I_{BE} = \frac{A_E J_S}{\beta_F} e^{v_e/n_f V_T} \tag{2.21}$$

where

A_E is the aforementioned emitter–base junction area
J_S is the density of transistor saturation current
β_F is the forward short-circuit current transfer ratio
n_f is the injection coefficient of the emitter–base junction
v_e is the internal junction voltage serving to forward bias the emitter–base junction

and

$$V_T = \frac{kT_j}{q} \tag{2.22}$$

is the Boltzmann voltage. In the last expression, k is Boltzmann's constant [1.38 (10^{-23}) J/K], T_j is the absolute temperature of the emitter–base junction, and q is the magnitude of electron charge [1.6 (10^{-19}) C].

The current I_{CC} is derived from [5]

$$I_{CC} = A_E J_S e^{v_e/n_f V_T} \left(1 - \frac{I_{CC}}{I_{KF}}\right)\left(1 + \frac{v_x}{V_{AF}}\right) \tag{2.23}$$

where

I_{KF}, which is proportional to A_E, is the "forward knee current" of the transistor [6]
V_{AF}, which is independent of A_E, is the "forward Early voltage" [7]

Note that the base current i_b is the current I_{BE}, while the collector current i_c is I_{CC}. Thus, the "static common-emitter current" gain (often referred to as the "DC beta"), h_{FE}, of a BJT is

$$h_{FE} = \frac{i_c}{i_b} = \frac{I_{CC}}{I_{BE}} = \beta_F \left(1 - \frac{i_c}{I_{KF}}\right)\left(1 + \frac{v_x}{V_{AF}}\right) \tag{2.24}$$

which is functionally dependent on both the collector current and the intrinsic collector–emitter voltage.

Unlike the base, collector, and emitter resistances, the resistance r_π in the small-signal model of Figure 2.16a is not an ohmic branch element. It is a mathematical resistance that arises from the Taylor series expansion of the current I_{BE} about the "quiescent-operating point," or "Q-point" of the transistor. In particular, r_π, which is known as the "emitter–base junction diffusion resistance," derives from

$$\frac{1}{r_\pi} = \left.\frac{\partial I_{BE}}{\partial v_e}\right|_Q \tag{2.25}$$

where it is understood that the indicated derivative is evaluated at the Q-point of the device. This Q-point is unambiguously defined by the "zero signal, or static," values of the base current I_{BQ}, the collector current I_{CQ}, and the internal collector–emitter voltage V_{XQ}. Using Equations 2.21 and 2.24, and the fact that $i_b + I_{BE}$,

$$r_\pi = \frac{h_{FE} n_f V_T}{I_{CQ}} \tag{2.26}$$

The inverse dependence of r_π on quiescent collector current renders r_π large at low collector current biases.

Similarly, r_o, the "forward Early resistance," derives from

$$\frac{1}{r_o} = \left.\frac{\partial I_{CC}}{\partial v_x}\right|_Q \tag{2.27}$$

It can be shown that

$$r_o = \frac{V_{XQ} + V_{AF}}{I_{CQ}\left(1 - \frac{I_{CQ}}{I_{KF}}\right)} \tag{2.28}$$

Like r_π, r_o is also large for low-level biasing.

Finally, the parameter β, which is the "low-frequency small-signal common-emitter short-circuit current gain" (often more simply referred to as the "AC beta") of the transistor, is

$$\beta = g_m r_\pi \tag{2.29}$$

where g_m, the "forward transconductance" of a BJT is

$$g_m = \left.\frac{\partial I_{CC}}{\partial v_e}\right|_Q \tag{2.30}$$

From Equations 2.23, 2.24, 2.26, and 2.29,

$$\beta = h_{FE}\left(1 - \frac{I_{CQ}}{I_{KF}}\right) \tag{2.31}$$

To the extent that $I_{CQ} \ll I_{KF}$, β is nominally independent of both Q-point collector current and emitter–base junction injection area.

2.2.3 Single-Input–Single-Output Canonic Cells

2.2.3.1 Diode-Connected Transistor

The simplest of the single-input–single-output, or "single-ended" canonic cells for linear bipolar circuits is the "diode-connected transistor" offered in Figure 2.17a. This transistor connection emulates the volt–ampere characteristics of a conventional PN junction diode. It can therefore be used in rectifier, voltage regulator, dc level shifting, and other applications that exploit conventional diodes. But unlike a conventional PN junction diode, the diode-connected transistor proves especially useful in current mirror biasing schemes. These and other similar circuits require that the base–emitter terminal voltage, v, of the diode track the base–emitter terminal voltage of a second, presumably identical transistor, over wide variations in junction-operating temperatures.

If the voltages dropped across the internal base, collector, and emitter resistances are small, the intrinsic emitter–base junction voltage, v_e, is approximately the indicated terminal voltage, v. Moreover, the intrinsic base–collector junction voltage, v_c, is essentially zero. It follows that for $v > v_\gamma$, the transistor in the diode connection operates in its linear regime.

In the subject diagram, the terminal voltage v is depicted as a superposition of a static voltage, V_Q, and a signal component, v_s. The resultant diode current, i, is a superposition of a quiescent current, I_Q, and a signal current, i_s. The quiescent components of diode voltage and current arise from static power supplied to the diode circuit to ensure that the diode-connected device operates in its linear regime. On the other hand, the signal components are established by a time-varying signal applied to the input port of the circuit in which the diode-connected transistor is embedded. In order to achieve reasonably linear processing of the applied input signal, the value of V_Q must be such as to ensure that $v = V_Q + v_s > v_\gamma$ for all values of the time-varying signal voltage v_s. Since v_s can be positive or negative at any instant of time, the requirement $V_Q + v_s > v_\gamma$ mandates that the amplitude of v_s be sufficiently small.

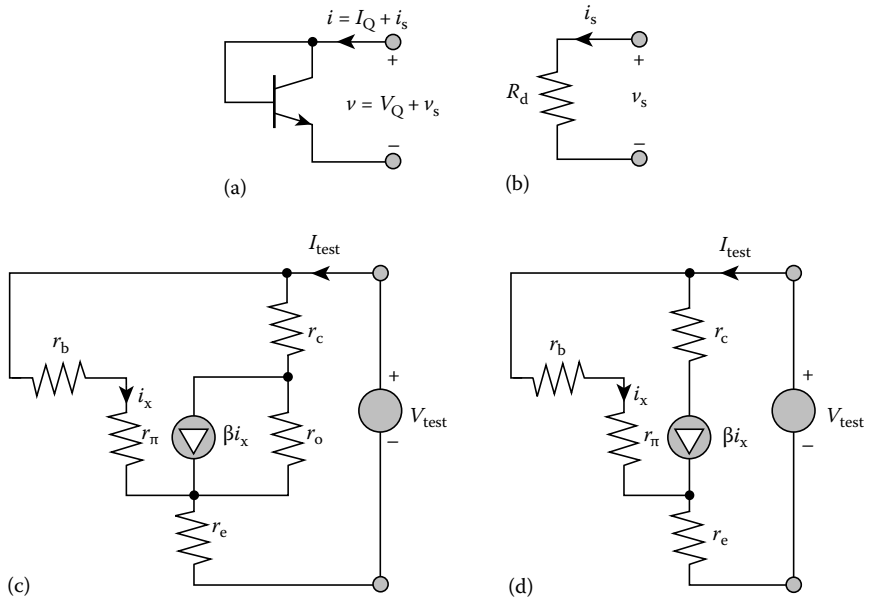


FIGURE 2.17 (a) Diode-connected BJT. (b) The small-signal equivalent circuit of the diode in (a). (c) Low-frequency, small-signal model of the diode-connected transistor in (a). The ratio V_{test}/I_{test} is the small-signal resistance R_d presented at the terminals of the diode-connected transistor. (d) The model in (c) approximated for the case of very large Early resistance.

The immediate impact of the small-signal condition corresponding to the linearity requirement $V_Q + v_s > v_\gamma$ is that the small-signal volt–ampere characteristics of the diode are linear. And since the diode is a two-terminal element, these characteristics can be modeled at low-signal frequencies by a simple resistance, say R_d, as suggested by the single-element macromodel offered in Figure 2.17b. The resistance in the latter figure can be determined by using the small-signal transistor model of Figure 2.16a to construct the small-signal equivalent circuit of the diode-connected transistor shown in Figure 2.17c. In this figure, the ratio of the test voltage, V_{test}, to the test current, I_{test}, is the desired resistance, R_d. A straightforward KVL analysis confirms that

$$R_d = \frac{V_{test}}{I_{test}} = r_e + \frac{(r_o + r_c) \| (r_b + r_\pi)}{1 + \frac{\beta r_o}{r_o + r_c + r_b + r_\pi}} \tag{2.32}$$

Typically, r_o is 25 kΩ or larger, r_c is smaller than 75 Ω, r_b is of the order of 100 Ω, and r_π is in the range of 1 kΩ for a minimal geometry device. It follows that $r_o \gg (r_c + r_b + r_\pi)$, and R_d can be approximated as

$$R_d \approx r_e + \frac{r_b + r_\pi}{\beta + 1} \tag{2.33}$$

Note that this terminal resistance is of the order of the low tens of ohms. For example, if $r_b = 100$ Ω, $r_\pi = 1.2$ kΩ, $r_e = 1$ Ω, and $\beta = 100$, $R_d = 13.9$ Ω. It is instructive to note that the approximation, $r_o \gg (r_c + r_b + r_\pi)$, collapses the model in Figure 2.17c to the structure in Figure 2.17d, from which Equation 2.33 follows immediately.

A variation of the diode scheme is the so-called V_{BE} "multiplier" depicted in Figure 2.18a. This circuit finds extensive use in regulator and level shifting applications that require either a series interconnection of more than one diode or a circuit branch voltage drop whose requisite value is a nonintegral multiple of the base–emitter terminal voltage of a single diode.

The circuit under consideration establishes a static terminal voltage, V_Q, whose value is a designable multiple of the static base–emitter terminal voltage, V_{BEQ}. To confirm this contention, observe that for static operating conditions, V_Q is

$$V_Q = R_Y(I_Q - I_{CQ}) + V_{BEQ} \tag{2.34}$$

where the voltage component V_{BEQ} of the net base–emitter terminal voltage v_{be} is

$$V_{BEQ} = R_X \left[I_Q - \left(\frac{h_{FE} + 1}{h_{FE}} \right) I_{CQ} \right] \tag{2.35}$$

The current, I_{CQ}, is the static component of the net collector current, i_c, and h_{FE} is the collector current to base current transfer ratio defined by Equation 2.24. An elimination of I_{CQ} from the foregoing two expressions leads to

$$V_Q = \left(1 + \frac{\alpha_{FE} R_Y}{R_X} \right) V_{BEQ} + \left(\frac{R_Y}{h_{FE} + 1} \right) I_Q \tag{2.36}$$

where

$$\alpha_{FE} = \frac{h_{FE}}{h_{FE} + 1} \tag{2.37}$$

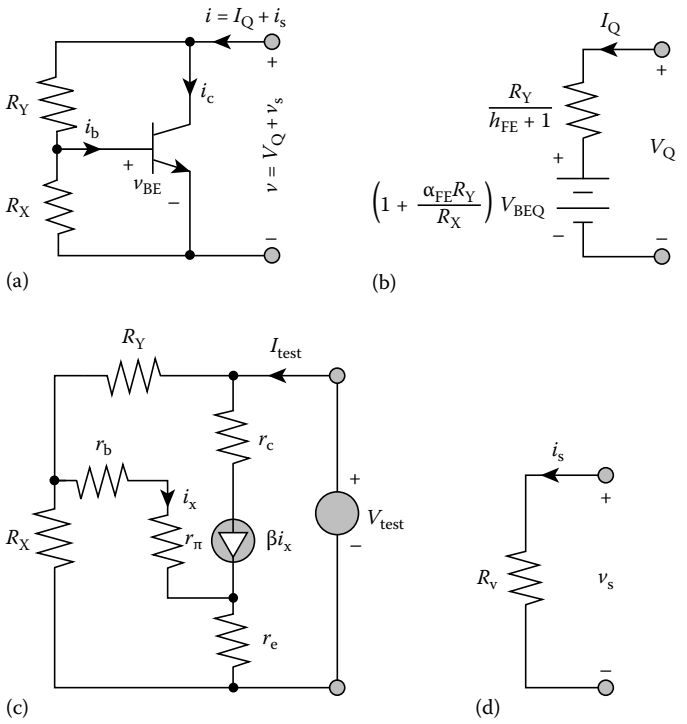


FIGURE 2.18 (a) Schematic diagram of V_{BE} multiplier. (b) DC macromodel of the multiplier in (a). (c) Low-frequency small-signal equivalent circuit of the Y_{BE} multiplier. (d) The small-signal equivalent resistance at the terminals of the V_{BE} multiplier.

is known as the ''static common-base current gain'' (often referred to as the ''DC alpha'') of a BJT. Equation 2.36 suggests that the static electrical behavior of the V_{BE} multiplier approximates a battery, whose voltage is controllable by the resistive ratio R_Y/R_X. The internal resistance of this effective battery is inversely dependent on $(h_{FE} + 1)$, and is therefore small. The macromodel in Figure 2.18b reflects the foregoing electrical interpretation. Note the for $R_Y = 0$ and R_X infinitely large, the circuit in Figure 2.18a collapses to the diode-connected transistor of Figure 2.17a, and V_Q understandably reduces to V_{BEQ}, the quiescent base–emitter terminal voltage of a diode-connected transistor.

For $V_Q + v_s > v_\gamma$, the transistor in the V_{BE} multiplier operates linearly. Accordingly, the pertinent small-signal terminal characteristics emulate a resistance, say R_V, which can be determined by applying the model of Figure 2.16a to the circuit in Figure 2.18a. The resultant equivalent circuit, simplified to reflect the realistic assumption of large r_o, is shown in Figure 2.18c while Figure 2.18d postulates the small-signal macromodel. An analysis of the circuit in Figure 2.18c reveals that

$$R_v \approx R_X \| R_d + R_Y \left[1 - \frac{\alpha R_X}{R_X + R_d}\right] \tag{2.38}$$

where

$$\alpha = \frac{\beta}{\beta + 1} \tag{2.39}$$

is the "low-frequency, small-signal, common-base, short-circuit current gain" (more simply referred to as the "ac alpha") of the transistor, and R_d is the resistance given by Equation 2.32. For $R_Y = 0$, $R_X = \infty$ reduces R_v to the expected result, $R_v \approx R_d$. Note further that for $\beta \gg 1$ (which makes $\alpha \approx 1$) and $R_d \gg R_X$, R_v is essentially the small-signal resistance presented at the terminals of a diode-connected transistor.

2.2.3.2 Common-Emitter Amplifier

The most commonly used single-ended canonic gain cell is the "common-emitter amplifier," whose NPN and PNP AC schematic diagrams are shown in Figure 2.19a and b, respectively. The AC schematic diagram delineates only the signal paths of a circuit. Thus, the biasing subcircuits required for linear operation of the transistors are not shown, thereby affording topological and analytical simplification. This simplification is accomplished without loss of engineering generality, for the results produced by an analysis of the AC schematic diagram reveal all salient performance traits of the common-emitter configuration.

The common-emitter amplifier is distinguished by the facts that signal is applied to the base of the transistor, and the resultant response is extracted as either the voltage, V_{OS}, or the current, I_{OS}, at the collector port. The effective load resistance terminating the collector to ground is indicated as R_{LT}, while the signal source is represented as a traditional Thévenin equivalent circuit. Alternatively, a Norton representation of the input source can be used, with the understanding that the Norton equivalent signal current, say I_{ST}, is simply the ratio of the Thévenin signal voltage, V_{ST}, to the Thévenin source resistance, R_{ST}.

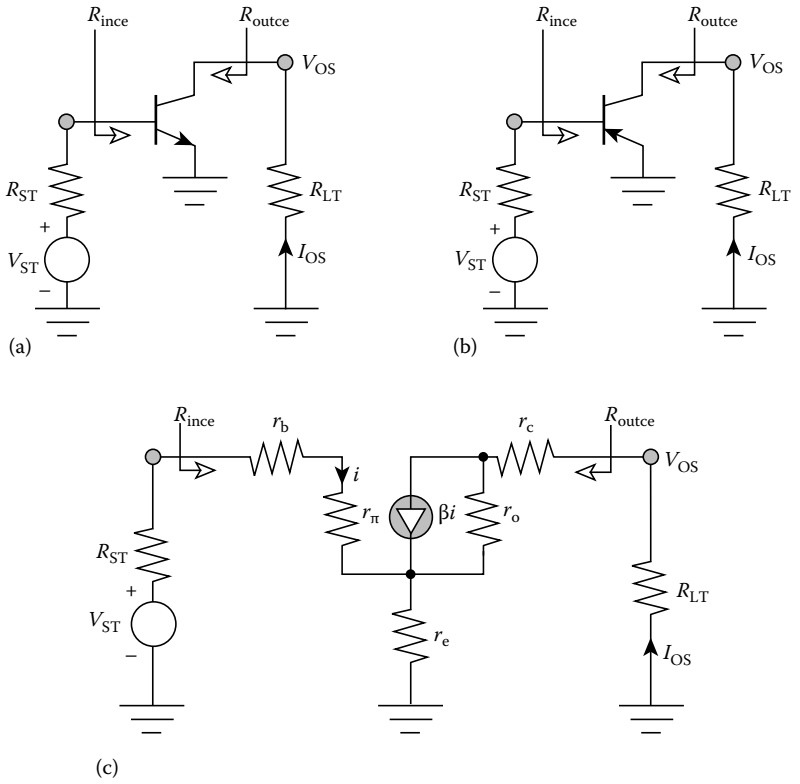


FIGURE 2.19 (a) AC schematic diagram of an NPN common-emitter amplifier. (b) AC schematic diagram of a PNP common-emitter amplifier. (c) Small-signal, low-frequency equivalent circuit of the common-emitter amplifier.

The common-emitter amplifier is capable of large magnitudes of voltage and current gains, moderately large input resistance, and very large driving point output resistance. An analytical confirmation of these contentions begins by drawing the small-signal equivalent circuit of the amplifier. This structure is given in Figure 2.19c and is valid for either the NPN or the PNP versions of the amplifier. An analysis of the small-signal model yields a voltage gain,

$A_{vce} = V_{OS}/V_{ST}$, of

$$A_{vce} = -\left\{\frac{(\beta - (r_e/r_o))(r_o/(r_o + r_c + r_e + R_{LT}))R_{LT}}{R_{ST} + r_b + r_\pi + ((\beta r_o/(r_o + r_c + R_{LT})) + 1)[r_e \| (r_o + r_c + R_{LT})]}\right\} \tag{2.40}$$

This relationship can be simplified by exploiting the fact that the internal resistance r_e of a transistor is small. Thus, $\beta \gg r_e/r_o$ and $r_e \gg (r_o + r_c + R_{LT})$, thereby implying

$$A_{vce} \approx -\frac{\beta_{eff} R_{LT}}{R_{ST} + r_b + r_\pi + (\beta_{eff} + 1)r_e} \tag{2.41}$$

where

$$\beta_{eff} \triangleq \beta\left[\frac{r_o}{r_o + r_c + R_{LT}}\right] \tag{2.42}$$

is an attenuated version of the AC beta for the utilized transistor. This effective beta approximates β itself, since $r_o \gg (r_c + R_{LT})$ is typical.

In concert with earlier arguments, Equation 2.41 confirms a diminished magnitude of gain for large internal device resistances. Note also that phase inversion, as inferred by the negative sign in either Equation 2.40 or 2.41 prevails between the Thévenin source voltage, V_{ST}, and the voltage signal response, V_{OS}. Finally, observe that large magnitudes of voltage gain are possible in the common-emitter orientation when β_{eff} is sufficiently large.

The driving point input resistance, R_{ince}, of the common-emitter amplifier can be determined as the ratio V_x/I_x for the test structure depicted in Figure 2.20a. It is easily shown that

$$R_{ince} = r_b + r_\pi + (\beta_{eff} + 1)[r_e \| (r_o + r_c + R_{LT})] \tag{2.43}$$

Since $r_e \ll (r_o + r_c + R_{LT})$, Equation 2.43 collapses to

$$R_{ince} \approx r_b + r_\pi + (\beta_{eff} + 1)r_e \tag{2.44}$$

Similarly, the driving point output resistance, R_{outce}, is derived as the V_x/I_x ratio of the equivalent circuit offered in Figure 2.20b. In particular,

$$R_{outce} = r_c + r_e \| (r_\pi + r_b + R_{ST}) + \left(\frac{\beta r_e}{r_e + r_\pi + r_b + R_{ST}} + 1\right) r_o \tag{2.45}$$

Since the model resistance r_π varies inversely with collector bias current, R_{ince} is moderately large when the common-emitter transistor is biased at low currents. On the other hand, R_{outce} is very large since Equation 2.45 confirms $R_{outce} > r_o$.

When the foregoing results are simplified to reflect the practical special case of a very large forward Early resistance, r_o, the cumbersome small-signal equivalent circuit of Figure 2.19c reduces to a ''small-signal macromodel'' useful for design-oriented circuit analysis of multistage amplifiers. To this end, note that a large r_o produces a driving point common-emitter input resistance that is independent of the

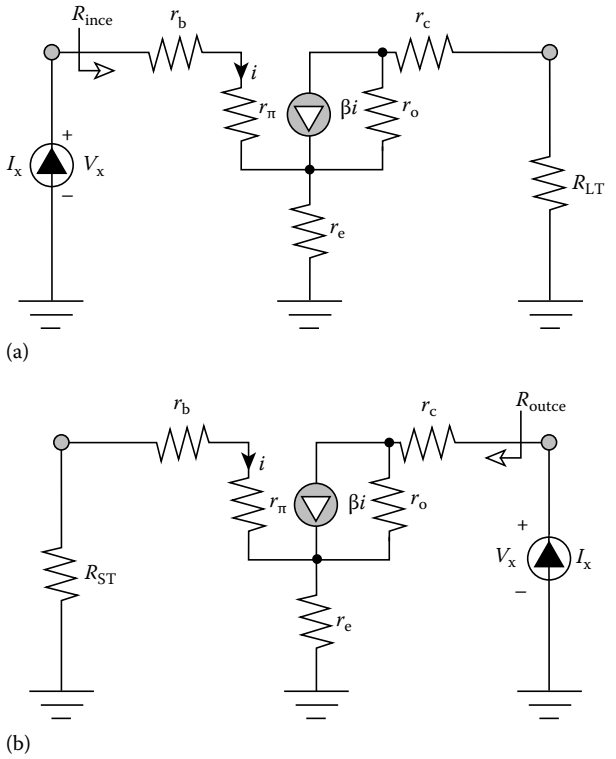


FIGURE 2.20 (a) Small-signal test structure used to determine the driving point input resistance of the common-emitter amplifier. (b) Small-signal test structure used to determine the driving point output resistance of the common-emitter amplifier.

terminating load resistance. Such independence implies no internal feedback from the output to input ports. It follows that the small-signal volt–ampere characteristics at the input port of a common-emitter amplifier can be modeled approximately by a simple resistance of value, R_{ince}, as defined by Equation 2.44. On the other hand, the large driving point output resistance R_{outce} suggests that a prudent output port model of a common-emitter stage is a Norton equivalent circuit. The Norton, or short-circuit, output current is proportional to the applied input signal voltage, V_{ST}, as depicted in Figure 2.21a. Alternatively, it can be expressed as a proportionality of the Norton input signal current, I_{ST}, as suggested in Figure 2.21b. In the former figure, the Norton current is

$$G_{fce}V_{ST} = \lim_{R_{LT}\to 0} I_{OS} = \lim_{R_{LT}\to 0}\left(-\frac{V_{OS}}{R_{LT}}\right) = \lim_{R_{LT}\to 0}\left(-\frac{A_{vce}V_{ST}}{R_{LT}}\right) \tag{2.46}$$

Subject to the assumption of large r_o,

$$G_{fce} = \lim_{R_{LT}\to 0}\left(-\frac{A_{vce}}{R_{LT}}\right) \approx \frac{\beta}{R_{ST} + r_b + r_\pi + (\beta+1)r_e} \tag{2.47}$$

Recalling Equation 2.29, this effective forward transconductance of the amplifier can be expressed in terms of the transconductance, g_m, of the transistor utilized in the amplifier. Specifically,

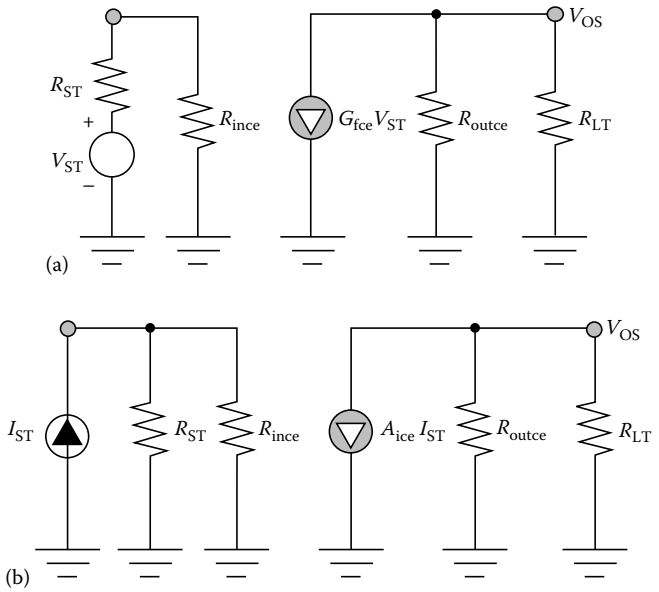


FIGURE 2.21 (a) Small-signal macromodel of a common-emitter amplifier in which the Norton output port circuit uses a voltage-controlled current source. (b) Small-signal macromodel of a common-emitter amplifier in which the Norton output port circuit uses a current-controlled current source.

$$G_{fce} \approx \frac{g_m}{1 = g_m r_e + ((r_e + r_b + R_{ST})/r_\pi)} \tag{2.48}$$

In the macromodel of Figure 2.21a, R_{outce} is very large by virtue of large r_o. Accordingly, the parallel combination of R_{outce} and R_{LT} is essentially R_{LT}, thereby implying an approximate common-emitter voltage gain of

$$A_{vce} \approx -G_{fce}R_{LT} \tag{2.49}$$

For the alternative macromodel in Figure 2.21b, the Norton current is

$$A_{ice}I_{ST} = A_{ice}\left(\frac{V_{ST}}{R_{ST}}\right) = \lim_{R_{LT}\to 0} I_{OS} = G_{fce}V_{ST} \tag{2.50}$$

Since $V_{ST} = R_{ST}I_{ST}$, it follows that A_{ice} is, for large r_o,

$$A_{ice} = G_{fce}R_{ST} \approx \beta\left[\frac{R_{ST}}{R_{ST} + r_b + r_\pi + (\beta + 1)r_e}\right] \tag{2.51}$$

Note that the Norton current proportionality A_{ice}, which is, in fact, the approximate ratio of the indicated output current, I_{OS}, to the Norton source current, I_{ST}, in the common-emitter configuration, is always smaller than β.

Example 2.1

Transistor Q_1 in the amplifier depicted in Figure 2.22a is fundamentally a common-emitter configuration since input signal is applied to its base terminal and the output voltage signal response is extracted at its

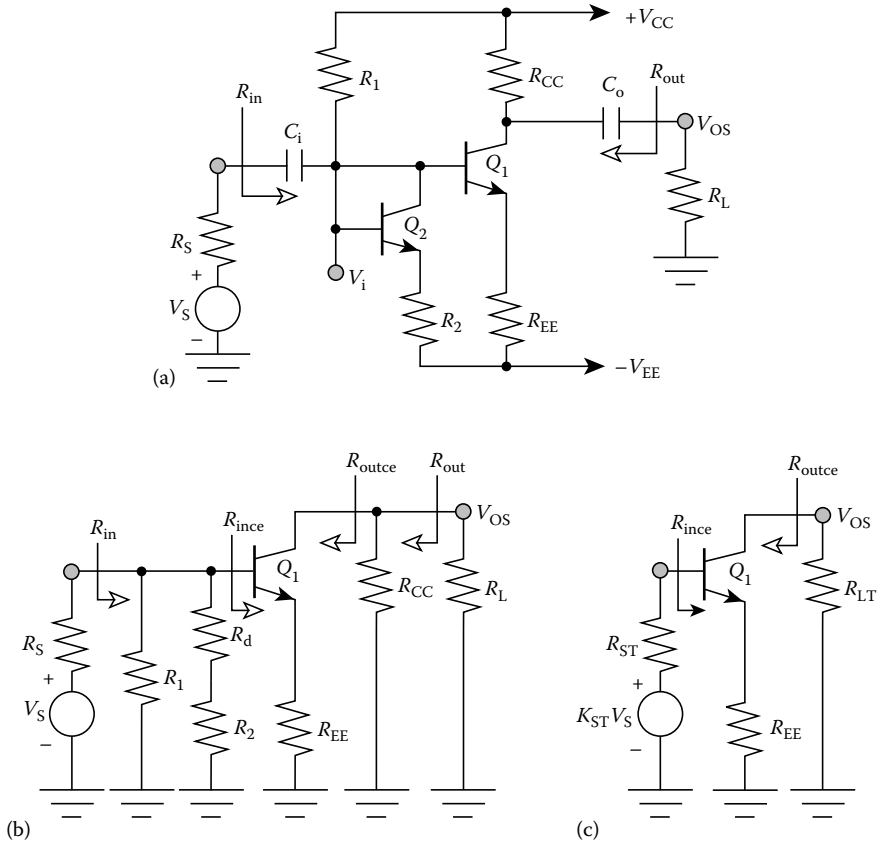


FIGURE 2.22 (a) Common-emitter amplifier with capacitively coupled input and output signal ports. (b) AC schematic diagram of the amplifier in (a). (c) Simplified AC schematic diagram of the amplifier in (a).

collector. The amplifier uses coupling capacitors C_i and C_o at its input and output ports. The input coupling capacitor, C_i, blocks the flow of static current in the source signal branch consisting of the series interconnection of the voltage, V_S, and the Thévenin source resistance, R_S. Accordingly, C_i, precludes R_S from affecting the biasing of both transistors used in the amplifier. Similarly, the output coupling capacitor, C_o, blocks the flow of static current in the external load resistance, R_L. Thus, both C_i and C_o can be viewed as open circuits for dc considerations. But simultaneously, these capacitors can be rendered transparent for AC situations by choosing them sufficiently large so that they emulate short circuits at the lowest frequency, say f_l, of signal-processing interest. In this problem, it is tacitly assumed that C_i and C_o, which are perfect DC open circuits, behave as good approximations of AC short circuits.

The subject amplifier utilizes a diode-connected transistor (Q_2) for temperature compensation of the static collector current conducted by transistor Q_1. For simplicity, assume that these two transistors have identical small-signal parameters of $r_b = 90\ \Omega$, $r_c = 55\ \Omega$, $r_\pi = 970\ \Omega$, $r_o = 42\ k\Omega$, and $\beta = 115$. Let the indicated circuit parameters be $R_1 = 2.2\ k\Omega$, $R_2 = 1.3\ k\Omega$, $R_{EE} = 75\ \Omega$, $R_{CC} = 3.9\ k\Omega$, $R_L = 1.0\ k\Omega$, and $R_S = 300\ \Omega$. Assuming that these circuit variables ensure linear operation of both devices, determine the small-signal voltage gain, $A_v = V_{OS}/V_S$, the driving point input resistance, R_{in}, and the driving point output resistance, R_{out}, of the amplifier. Finally, calculate the requisite minimum values of the input and output coupling capacitors, C_i and C_o, such that the lowest frequency f_l of interest is 500 Hz.

Solution

1. The first step of the solution process entails drawing the AC schematic diagram of the subject amplifier. By casting this diagram in the form of the canonic cell shown in Figure 2.19a, the gain and resistance expressions provided above can be exploited directly to assess the small-signal performance of the circuit at hand. Such a solution tack maximizes design-oriented understanding by avoiding the algebraic tedium implicit to an analysis of the entire small-signal equivalent circuit of the amplifier.

   To the foregoing end, observe that transistor Q_2 operates in its linear regime as a diode. It can therefore be viewed as the two terminal resistance R_d, given by Equation 2.32 or 2.33. Since the Early resistance is large, the latter expression can be used to arrive at $R_d = 10.64\ \Omega$.

   Since the power supply voltage, V_{EE}, is presumed ideal in the sense that it contains no signal component, the resultant series combination of R_d and R_2 returns base of transistor Q_1 to ground, as depicted in Figure 2.22b. Similarly, R_1 appears in shunt with the series interconnection of R_d and R_2, since V_{CC}, like V_{EE}, is also presumed to be an ideal constant (zero signal component) source of voltage. The AC schematic diagram of the input port is completed by noting that the AC short-circuit nature of the coupling capacitance C_i effectively connects the Thévenin representation of the signal source directly between the base of Q_1 and ground.

   At the output port of the amplifier, R_{CC} connects between the collector and ground since, as already exploited, V_{CC} is an AC short circuit. Moreover, the external load resistance, R_L, shunts R_{CC}, as shown in Figure 2.22b, because C_o behaves as an AC short circuit. The AC schematic diagram is completed by inserting the emitter degeneration resistance R_{EE} as a series element between ground and the emitter of transistor Q_1.
2. The diagram in Figure 2.22b can be straightforwardly collapsed to the simplified topology of Figure 2.22c. In the latter circuit, the effective load resistance, R_{LT}, is

$$R_{LT} = R_{CC} \| R_L = 795.9\ \Omega$$

   At the input port, the Thévenin resistance seen by the base of Q_1 is

$$R_{ST} = R_1 \| (R_d + R_2) \| R_S = 219.7\ \Omega$$

   while the corresponding Thévenin signal voltage can be expressed as $K_{ST} V_S$, where the voltage divider K_{ST} is

$$K_{ST} = \frac{R_1 \| (R_d + R_2)}{R_1 \| (R_d + R_2) + R_S} = 0.733$$

   The implication of this calculation is that, insofar as the active transistor Q_1 is concerned, the biasing resistances R_1 and R_2, cause a loss of more than 25% of the applied input signal.
3. The resultant AC schematic diagram in Figure 2.22c is virtually identical to the canonic topology in Figure 2.19a. Indeed, if the circuit resistance R_{EE} is absorbed into Q_1, where it appears in series with the internal emitter resistance r_e, the diagram is identical to the AC schematic diagram of the canonic common-emitter cell. Since $(r_e + R_{EE}) = 76.5\ \Omega$ is better than 560 times smaller than the resistance sum, $(r_o + r_c + R_{LT}) = 42.85\ \text{k}\Omega$, and $\beta = 115$ is more than 63,000 times larger than the resistance ratio $(r_e + R_{EE})/r_o = 0.00182$, the simplified expression in Equation 2.41 can be used to evaluate the voltage gain $V_{OS}/K_{ST}V_S$ in Figure 2.22c. From Equation 2.42, the effective small-signal beta is $\beta_{eff} = 112.7$. Then, with r_e replaced by $(r_e + R_{EE}) = 76.5\ \Omega$. Equation 2.41 gives

$$A_{vce} = \frac{V_{OS}}{K_{ST} V_S} = -8.99\ \text{V/V}$$

   It follows that the actual voltage gain of the amplifier in Figure 2.22a is

$$A_v = \frac{V_{OS}}{V_S} = K_{ST}A_{vce} = -6.59 \text{ V/V}$$

A better design, in the sense of achieving an adequate desensitization of circuit transfer characteristics with respect to parametric uncertainties, entails the use of a slightly larger emitter degeneration resistance R_{EE} selected to ensure that $(\beta_{eff}+1)(r_e+R_{EE}) \approx \beta_{eff} R_{EE} \gg (R_{ST}+r_b+r_\pi)$. For such a design, Equation 2.41 produces

$$A_v \approx -\frac{K_{ST}R_{LT}}{R_{EE}}$$

which is nominally independent of transistor parameters.

4. With r_e replaced by $(r_e+R_{EE}) = 76.5\ \Omega$, Equation 2.44 gives for the input resistance seen looking into the base of transistor Q_1 in Figure 2.22c, $R_{ince} = 9.76\ \text{k}\Omega$. It follows that the driving point input resistance seen by the source circuit in Figure 2.22b is

$$R_{in} = R_1 \| (R_d + R_2) \| R_{ince} = 757.6\ \Omega$$

5. The output resistance, R_{outce}, seen looking into the collector of transistor Q_1 in the diagram of Figure 2.22c is derived from Equation 2.45. With r_e replaced by $(r_e+R_{EE}) = 76.5\ \Omega$, $R_{outce} = 2.49\ \text{M}\Omega$. The resultant driving point output resistance seen by the load circuit in Figure 2.22b is

$$R_{out} = R_{CC} \| R_{outce} = 3894\ \Omega$$

The circuit output resistance is only a scant 6 Ω smaller than the collector biasing resistance R_{CC}, owning to the large value of R_{outce}. In turn, the value of the latter resistance is dominated by the last term on the right-hand side of Equation 2.45, which is proportional to the large forward Early resistance, r_o.

6. The input coupling capacitance, C_i, can be calculated with the help of the input port macromodel of Figure 2.23a. In this model, the subcircuit to the right of C_i in Figure 2.22a is replaced by its Thévenin equivalent circuit, which consists of the driving point input resistance, R_{in}, calculated previously. The voltage transfer function of this input port is

$$\frac{V_i(j\omega)}{V_S(j\omega)} = \left(\frac{R_{in}}{R_{in}+R_S}\right)\left[\frac{j\omega(R_{in}+R_S)C_i}{1+j\omega(R_{in}+R_S)C_i}\right]$$

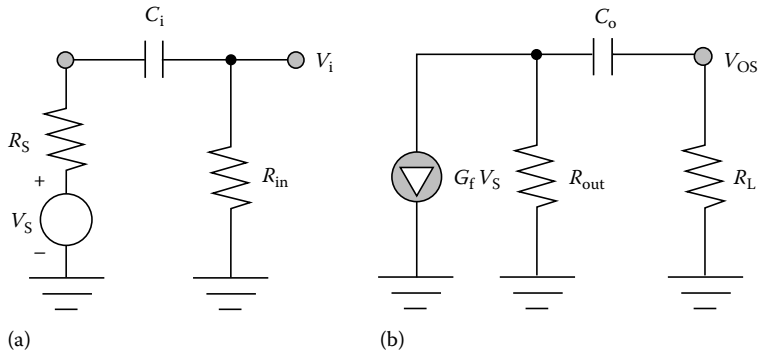


FIGURE 2.23 (a) Input port macromodel used in the calculation of the input coupling capacitor C_i. (b) Output port macromodel used to calculate the output coupling capacitor C_o.

An inspection of the foregoing relationship confirms that the dynamical effect of C_i on the voltage transfer function response is minimized if $\omega(R_{in} + R_s)C_i \gg 1$. At $\omega = 2\pi f_l$, the value of C_i that makes the left-hand side of this inequality equal to one is $C_i = 0.30$ μF. Observe that at $\omega_l(R_{in} + R_S)C_i = 1$,

$$\left|\frac{V_i(j\omega_1)}{V_S(j\omega_1)}\right| = \left(\frac{R_{in}}{R_{in} + R_S}\right)\left|\frac{j}{1+j}\right| = \frac{1}{\sqrt{2}}\left(\frac{R_{in}}{R_{in} + R_S}\right)$$

that is the magnitude of the input port voltage transfer function is a factor of the square root of two, or 3 dB, below the transfer function value realized at signal frequencies that are significantly higher than f_l. If this 3 dB attenuation is acceptable, $C_i = 0.30$ μF is appropriate to the design requirement.

7. The output coupling capacitance, C_o, is calculated analogously by exploiting the macromodel concepts overviewed in Figure 2.21a. To this end, the output port macromodel is offered in Figure 2.23b, where the effective forward transconductance, G_f, is such that $-G_f(R_{out}||R_L) = A_V$, as calculated in step 3. The voltage gain is seen to be

$$A_v(j\omega) = \frac{V_{OS}(j\omega)}{V_S(j\omega)} = -[G_f(R_{out}\|R_L)]\left[\frac{j\omega(R_{out} + R_L)C_o}{1 + j\omega(R_{out} + R_L)C_o}\right]$$

which has an algebraic form that is similar to the foregoing transfer relationship for the amplifier input port Thus, C_o is

$$C_o \geq \frac{1}{2\pi f_l(R_{out} + R_L)} = 0.065\ \mu F$$

Since the 0.30 μF input capacitor and the 0.065 μF output capacitor establish identical input and output port left-half-plane poles at the same frequency (f_1), the resultant attenuation at f_1 is actually larger than 3 dB. If this enhanced attenuation is unacceptable, the smaller of the two coupling capacitances can be made larger by a factor of three or so, thereby translating the associated pole frequency downward by a factor of three. In the present case, a plausible value of C_o is $C_o \geq (3)(0.065\ \mu F) = 0.2\ \mu F$.

It should be noted that the requisite two coupling capacitances are orders of magnitude too large for monolithic realization. Accordingly, if the subject amplifier is an integrated circuit, C_i and C_o are necessarily off-chip elements.

Example 2.2

When very large magnitudes of voltage gains are required, the output port of a common-emitter configuration can be terminated in an active load, as opposed to the passive resistive load encountered in the preceding example. Consider, for example, the complementary, NPN–PNP transistor amplifier whose schematic diagram appears in Figure 2.24a. The subcircuit containing transistors Q_1 and Q_2 is identical to that of the amplifier in Figure 2.22a. Indeed, for the purpose of this example, let the resistances, R_1, R_2, R_{EE}, and R_S, as well as small-signal parameters of transistors Q_1 and Q_2, remain at the values respectively stipulated for them in the preceding example. In the present diagram, the PNP transistor Q_3, along with its peripheral biasing resistances R_3, R_4, and R_5, supplants the resistance R_{CC} in the previously addressed common-emitter unit. Since no signal is applied to the base of Q_3, the subcircuit consisting of Q_3, R_3, R_4, and R_5 serves only to supply the appropriate biasing current to the collector of transistor Q_1. To the extent that this static current is invariant with temperature and the voltage signal response, V_{os}, established at the collector of Q_1, the Q_1 load circuit functions as a nominally constant current source. As a result, the effective load resistance, indicated as R_L in the subject figure,

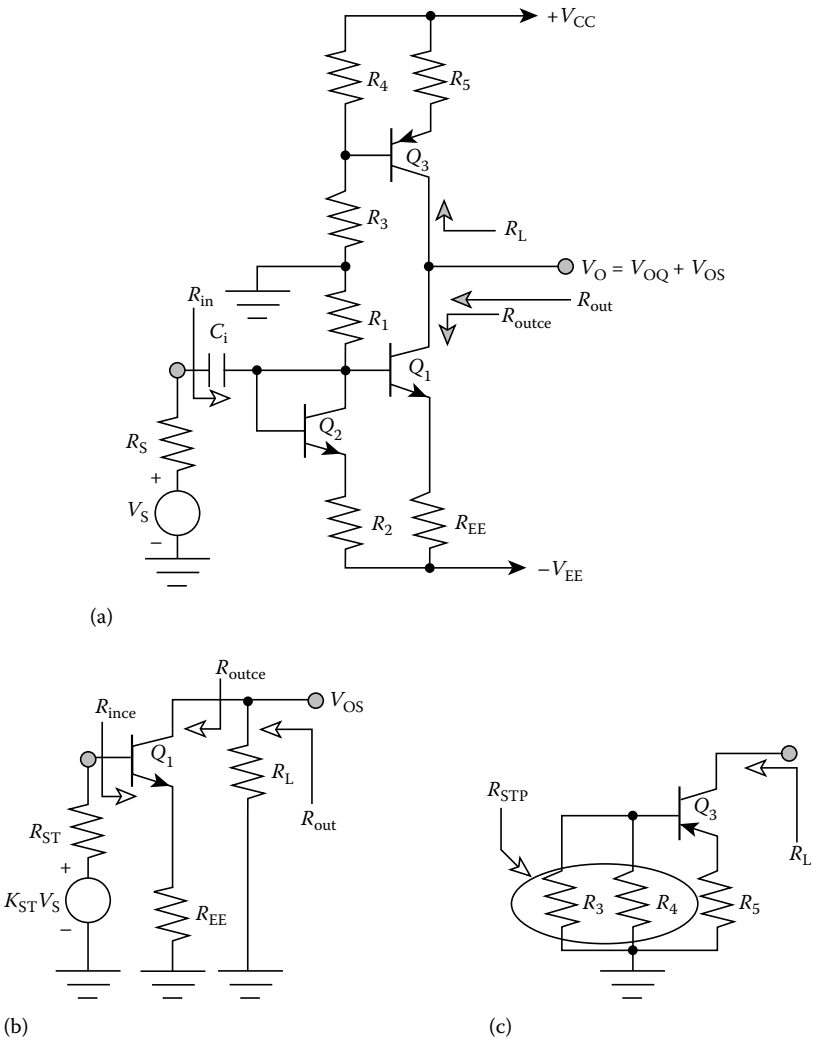


FIGURE 2.24 (a) Common-emitter amplifier with active current source load. (b) AC schematic diagram of the common-emitter unit. (c) AC schematic diagram of the active PNP transistor load.

seen by the collector of Q_1 is very large. In view of the absence of an external load appended to the output port, the resultant voltage gain of the amplifier is commensurately large.

Let $R_3 = 1.8$ kΩ, $R_4 = 3.3$ kΩ, and $R_5 = 100$ Ω. Moreover, let the small-signal parameters of the PNP transistor be $r_{bp} = 40$ Ω, $r_{cp} = 70$ Ω, $r_{ep} = 9$ Ω, $r_{\pi p} = 1100$ Ω, $r_{op} = 30$ kΩ, and $\beta_p = 60$. Assuming linear operation of all devices, determine the small-signal voltage gain $A_v = V_{OS}/V_S$, the driving point input resistance R_{in}, and the driving point output resistance R_{out} of the amplifier. As in the preceding example, the input coupling capacitance C_i can be presumed to act as an ac short circuit for the signal frequencies of interest.

Solution

1. The ac schematic diagram of the amplifier in Figure 2.24a is given in Figure 2.24b, where the PNP transistor load subcircuit is represented as an effective two terminal load resistance, R_L.

This representation is rendered possible by the fact that no signal is applied to the PNP load, which therefore acts only to supply biasing current to the collector of transistor Q_1. In the diagram, K_{ST}, R_{ST}, and R_{EE} (which effectively appears in series with r_e, the internal emitter resistance of Q_1) remain the same as in Example 2.1, namely, $K_{ST} = 0.733$, $R_{ST} = 219.7\ \Omega$, and $R_{EE} = 75\ \Omega$.

2. The AC schematic diagram of the Q_3 subcircuit alone appears in Figure 2.24c. A comparison of this figure with that shown in Figure 2.24b suggests that the subject diagram represents a PNP common-emitter amplifier under zero signal conditions. In particular, the Thévenin source resistance seen by the base of the PNP unit is $R_{STP} = R_3||R_4 = 1165\ \Omega$, while the emitter degeneration resistance of this subcircuit is $R_5 = 100\ \Omega$. It follows that the effective AC load resistance R_L terminating the collector port of Q_1 is the driving point output resistance of a common-emitter stage. With $r_c \triangleq r_{cp} = 70\ \Omega$, $r_b \triangleq r_{bp} = 40\ \Omega$, $r_e \triangleq (r_{ep} + R_5 = 109\ \Omega$, $r_\pi \triangleq r_{\pi p} = 1.1\ k\Omega$, $R_{ST} \triangleq R_{STP} = 1165\ \Omega$, $r_o \triangleq r_{op} = 30\ k\Omega$, and $\beta \triangleq \beta_p = 60$, (Equation 2.45) yields $R_{outce} \triangleq R_L = 111.5\ k\Omega$.
3. The voltage gain, input resistance, and output resistance of the actively loaded common-emitter amplifier can now be computed. For $r_c \triangleq r_{cn} = 55\ \Omega$, $r_b \triangleq r_{bn} = 90\ \Omega$, $r_e \triangleq (r_{en} + R_{EE} = 76.5\ \Omega$, $r_\pi \triangleq r_{\pi n} = 970\ k\Omega$, $R_{ST} \triangleq R_{STN} = 219.7\ \Omega$, $r_o \triangleq r_{on} = 42\ k\Omega$, $R_{LT} \triangleq R_L = 111.5\ k\omega$, and $\beta \triangleq \beta_n = 115$, Equation 2.42 gives an effective NPN transistor beta of $\beta_{eff} = 31.46$, and Equation 2.41 yields a voltage gain of $A_{vce} = V_{OS}/K_{ST}V_{ST} = -931.9$. It follows that the small-signal voltage gain of the stage at hand is $A_v = V_{OS}/V_S = K_{ST}A_{vce} = -682.6\ V/V$. It should be noted that this voltage gain is the ratio of only the signal component, V_{OS}, of the net output voltage, V_O (which contains a quiescent component of V_{OQ}) to the source signal voltage, V_S.
4. From Equation 2.43, the driving point input resistance seen looking into the base of transistor Q_1 in Figure 2.22b is $R_{ince} = 3.54\ k\Omega$. Then,

$$R_{in} = R_1||(R_d + R_2)||R_{ince} = 666.7\ \Omega$$

This input resistance differs slightly from the corresponding calculation in the preceding example owing to the reduction in the effective forward AC beta caused by the large active load resistance.

5. The resistance R_{outce} seen looking into the collector of transistor Q_1 remains the same as calculated in Example 2.1, namely, $R_{outce} = 2.49\ M\Omega$. It follows that the driving point output resistance of the amplifier under investigation is

$$R_{out} = R_L||R_{outce} = 106.7\ k\Omega$$

This large output resistance means that the actively loaded common-emitter configuration is a relatively poor voltage amplifier. In particular, an output buffer is mandated to couple virtually any practical external load resistance to the amplifier output port. In addition to reducing the output resistance, such a properly designed and implemented output buffer can reliably establish and stabilize the quiescent output voltage, V_{OQ}.

2.2.3.3 Common-Base Amplifier

The second of the canonic linear bipolar gain cells in the ''common-base amplifier,'' whose NPN and PNP AC schematic diagrams and corresponding small-signal equivalent circuit appear in Figure 2.25a and b, respectively. As it is confirmed below, the input resistance, R_{incb}, of this stage is very small and the output resistance, R_{outcb}, is very large. Accordingly, the common-base unit comprises a relatively poor voltage amplifier in the sense that its voltage gain, though potentially large, is a sensitive function of both the Thévenin source resistance R_{ST} and the Thévenin load resistance R_{LT}.

Although the common-base amplifier is not well suited for general voltage gain applications, it is an excellent ''current buffer,'' which is ideally characterized by zero input resistance, infinitely large output resistance, and unity current gain. When used for current buffering purposes, the common-base amplifier

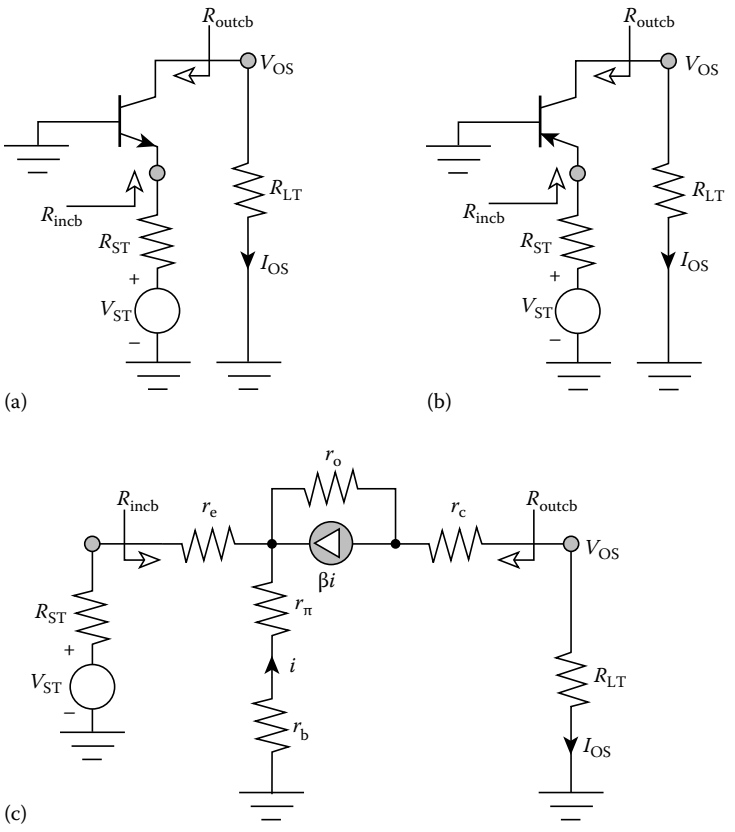


FIGURE 2.25 (a) AC schematic diagram of an NPN common-base amplifier. (b) AC schematic diagram of a PNP common-base amplifier. (c) Small-signal, low-frequency equivalent circuit of the common-base amplifier.

rarely appears as a stand alone single-stage amplifier, since signal excitations, particularly at the input and output ports of an electronic system are invariably formatted as voltages. Instead, it is invariably used in conjunction with an input voltage to current converter and/or an output current to voltage converter to achieve desired system performance characteristics.

The small-signal analysis of the common-base stage is considerably simplified if the assumption of large r_o is exploited at the outset. To this end, the equivalent circuit shown in Figure 2.25b reduces to the structure of Figure 2.26a. In the latter equivalent circuit, observe a signal emitter current i_{es} that relates to the indicated signal base current i in accordance with the Kirchhoff's current law constraint

$$i_{es} = -(\beta + 1)i \tag{2.52}$$

The signal component of the output current I_{OS} is therefore expressible as

$$I_{OS} = -\beta_i = \left(\frac{\beta}{\beta + 1}\right) i_{es} = \alpha i_{es} \tag{2.53}$$

where Equations 2.52 and 2.53 are used. The last result suggests the alternative model in Figure 2.26b, which is slightly more convenient version of the model in Figure 2.26a in that the current-controlled current source, αi_{es}, is dependent on the signal input port current i_{es}, as opposed to the signal current, i, that flows in the grounded base lead.

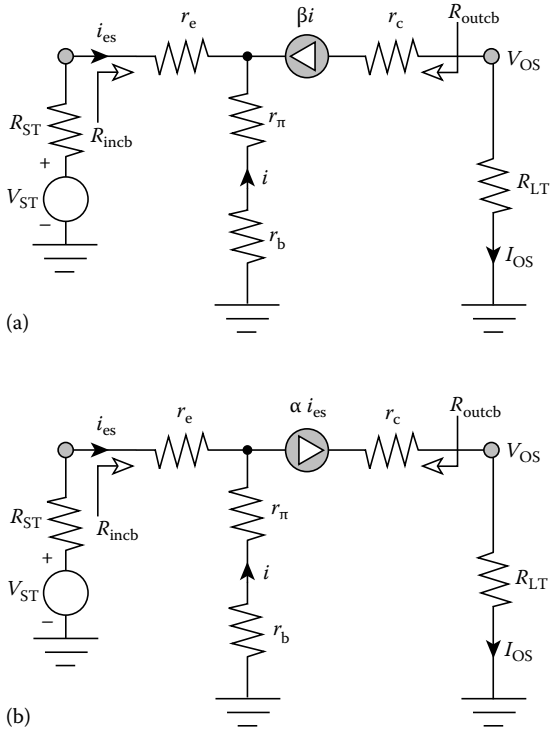


FIGURE 2.26 (a) The equivalent circuit of Figure 2.25c, simplified for the case of very large Early resistance r_o. (b) Modification of the circuit in (a) in which the current-controlled current source is rendered dependent on the input signal current i_{es}.

By inspection of the equivalent circuit in Figure 2.26b, the small-signal voltage gain A_{vcb} of the common-base cell is

$$A_{vcb} = \frac{\alpha R_{LT}}{R_{ST} + r_e + (1-\alpha)(r_\pi + r_b)} = \frac{\alpha R_{LT}}{R_{ST} + R_d} \tag{2.54}$$

where Equation 2.39 is used once again and R_d is the diode resistance defined by Equation 2.33. In contrast to the common-emitter cell, the common-base stage has no voltage gain phase inversion. But like the common-emitter configuration, the common-base voltage gain is directly proportional to the effective load resistance. It is also almost inversely proportional to the effective source resistance, given that the diode resistance R_d is small.

Although the voltage gain is vulnerable to uncertainties in the terminating load and source resistances, the common-base current gain, A_{icb}, is virtually independent of R_{LT} and R_{ST}. This contention follows from the fact that A_{icb}, which is the ratio of I_{OS} to the Norton equivalent source current, V_{ST}/R_{ST}, is

$$A_{icb} = \left(\frac{R_{ST}}{R_{LT}}\right) A_{vcb} = \frac{\alpha R_{ST}}{R_{ST} + R_d} \tag{2.55}$$

which is independent of R_{LT} (to the extent that the Early resistance r_o can indeed be ignored). Since the signal in a current-drive amplifier is likely to have a large source resistance, $R_{ST} \gg R_d$, which

implies $A_{icb} \approx \alpha$, independent of R_{LT} and R_{ST}. Note that this approximate current gain is essentially unity, since α as introduced by Equation 2.39 approaches one for the typically encountered circumstances of large β.

The input and output resistances of the common-base amplifier follow immediately from an analysis of the model in Figure 2.26b. In particular, the driving point input resistance, R_{incb}, is

$$R_{incb} = \left(\frac{V_{ST}}{i_{es}}\right)\Bigg|_{R_{ST}=0} = r_e + (1-\alpha)(r_\pi + r_b) = R_d \tag{2.56}$$

where the numerical value is of the order of only a few tens of ohms. On the other hand, the driving point output resistance, R_{outcb}, is infinitely large since $V_{ST} = 0$ constrains i_{es}, and thus αi_{es}, to zero. In turn, $\alpha i_{es}, = 0$ means that R_{LT} in Figure 2.26b faces an open circuit, whence $R_{outcb} = \infty$.

To the extent that the common-base amplifier is excited from a signal current source and that the forward Early resistance of the utilized transistor is very large, the common-base amplifier is seen to have almost unity current gain, very low input resistance, and infinitely large output resistance. Its transfer characteristics therefore approximate those of an ideal current buffer. Of course, the finite nature of the forward Early resistance renders the observable driving point output resistance of a common-base cell large, but nonetheless finite. The actual output resistance can be determined as the V_x to I_x ratio in the ac schematic diagram of Figure 2.27a. The requisite analysis is algebraically cumbersome owing to the presence of r_o in shunt with the current-controlled current source in the equivalent circuit of Figure 2.25c. Fortunately, however, an actual circuit analysis can be circumvented by a proper interpretation of cognate common-emitter results formulated earlier.

In order to demonstrate the foregoing contention, consider Figure 2.27b, which depicts the AC schematic diagram for determining the driving point output resistance R_{outce} of a common-emitter amplifier. The only difference between the two AC schematic diagrams in Figure 2.27 is the topological placement of the effective source resistance, R_{ST}. In the common-base stage, this source resistance is in series with the emitter of a transistor with a base that is grounded. On the other hand, R_{ST} appears in the common-emitter configuration as an element in series with the base of a transistor whose emitter is ground. It follows that Equation 2.45 can be used to deduce an expression for R_{outcb}, provided that in Equation 2.45 R_{ST} is set to zero and r_e is replaced by $(r_e + R_{ST})$.

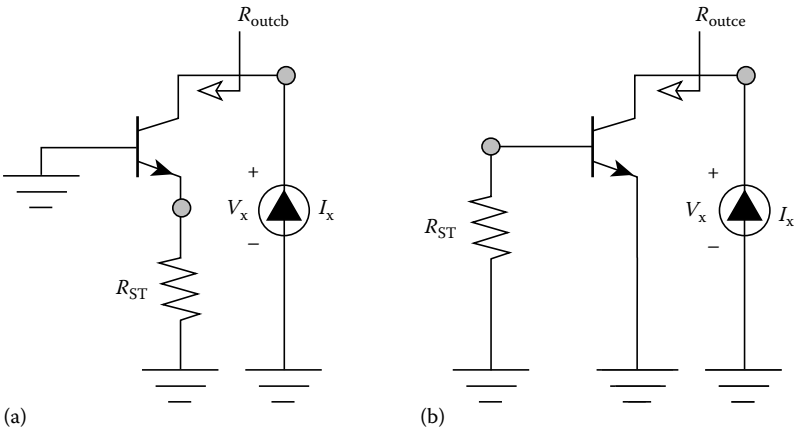


FIGURE 2.27 (a) AC schematic diagram appropriate to the computation of the driving point output resistance of a common-base amplifier. (b) AC schematic diagram pertinent to computing the driving point output resistance of a common-emitter amplifier.

The result is

$$R_{\text{outcb}} = r_c + (r_e + R_{ST}) \| (r_\pi + r_b) + \left[\frac{\beta(r_e + R_{ST})}{r_e + R_{ST} + r_\pi + r_b} + 1 \right] r_o \tag{2.57}$$

For large R_{ST} and large r_o, Equation 2.47 reduces to

$$R_{\text{outcb}} \approx (\beta + 1) r_o \tag{2.58}$$

which is an extremely large output resistance.

The common-base stage is generally used in conjunction with a common-emitter amplifier to form the "common-emitter-common-base cascode," whose schematic diagram is shown in Figure 2.28. In this application, the common-emitter stage formed by transistor Q_1, the emitter degeneration resistance, R_{EE}, and the biasing elements, R_1 and R_2, serves as a transconductor that converts the input signal voltage, V_S, to a collector current whose signal component is i_{1s}. Note that such conversion is encouraged by the fact that the effective load resistance, R_{Leff}, terminating the collector of Q_1, is the presumably low input resistance of the common-base stage formed by transistor Q_2 and the biasing resistances, R_3 and R_4. Since the current gain of a common-base stage is essentially unity, Q_2 translates the signal current in its emitter to an almost identical signal current flowing through the collector load resistance, R_L. The latter element acts as a current to voltage convert to establish the signal component, V_{OS}, of the net output voltage V_O.

The analysis of the common-emitter–common-base cascode begins by representing the collector port of the common-emitter configuration by its Norton equivalent circuit. Assuming that the input coupling

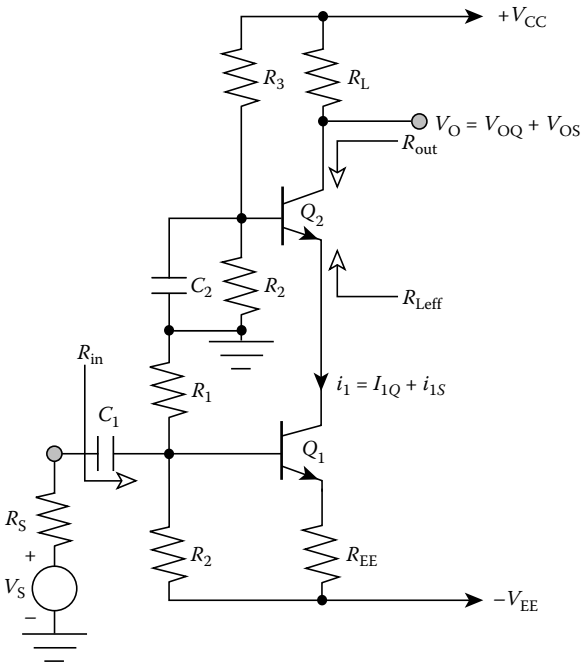


FIGURE 2.28 Schematic diagram of a common-emitter–common-base cascode. The common-emitter stage formed by transistor Q_1 and its peripheral elements acts as a voltage to current converter. Transistor Q_2 and its associated biasing elements function as a current amplifier, while the load resistance, R_L, acts as a current to voltage converter.

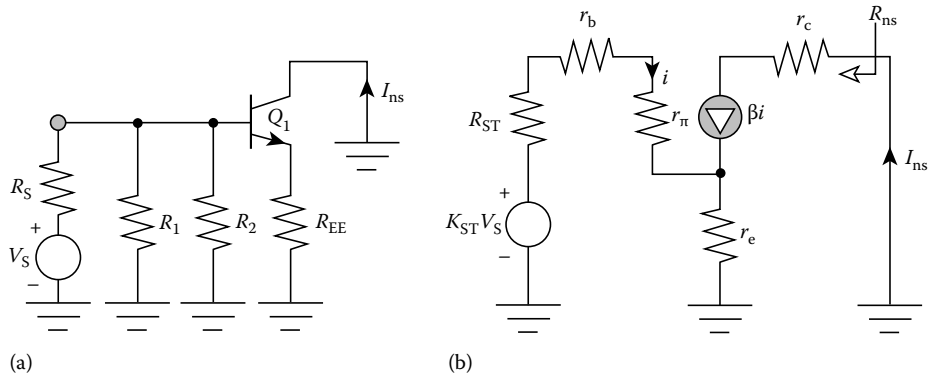


FIGURE 2.29 (a) AC schematic diagram used to calculate the Norton equivalent output circuit of the common-emitter subcircuit in the cascode configuration of Figure 2.28. (b) Small-signal model of the AC circuit in (a).

capacitor, C_1, is sufficiently large to enable its replacement by a short circuit over the signal frequency range of interest, the pertinent AC schematic diagram is the circuit in Figure 2.29a, where I_{ns} symbolizes the Norton, or short-circuit signal current conducted by the collector of transistor Q_1. The corresponding small-signal equivalent circuit appears in Figure 2.29b, where the Early resistance is tacitly ignored, the effective source resistance, R_{ST}, seen by the base of Q_1 is

$$R_{ST} = R_S \| R_1 \| R_2 \tag{2.59}$$

and the voltage divider K_{ST} is

$$K_{ST} = \frac{R_1 \| R_2}{R_1 \| R_2 + R_S} \tag{2.60}$$

Using the model in Figure 2.29b, it is a simple matter to show that

$$I_{ns} = \beta i = G_{ns} V_S \tag{2.61}$$

where G_{ns}, which can be termed the "Norton transconductance" of the common-emitter stage, is

$$G_{ns} = \frac{\beta K_{ST}}{R_{ST} + r_b + r_\pi + (\beta + 1) r_e} \tag{2.62}$$

The Norton output resistance R_{ns} is infinitely large by virtue of the assumption of infinitely large Early resistance.

The foregoing results permit drawing the AC schematic diagram of the common-base component of the common-emitter–common-base cascode in the topological form depicted in Figure 2.30a. From the corresponding small-signal equivalent circuit in Figure 2.30b, which assumes that the capacitor, C_2, behaves as an AC short circuit and which once again ignores transistor Early resistance, the voltage gain, say A_v, follows immediately as

$$A_v = \frac{V_{OS}}{V_S} = -\alpha G_{ns} R_L = -\frac{\alpha \beta K_{ST} R_L}{R_{ST} + r_b + r_\pi + (\beta + 1) r_e} \tag{2.63}$$

The driving point output resistance R_{out}, like the Norton output resistance of the common-emitter stage, is infinitely large. In fact, R_{out} is a good approximation of infinity. Its numerical value

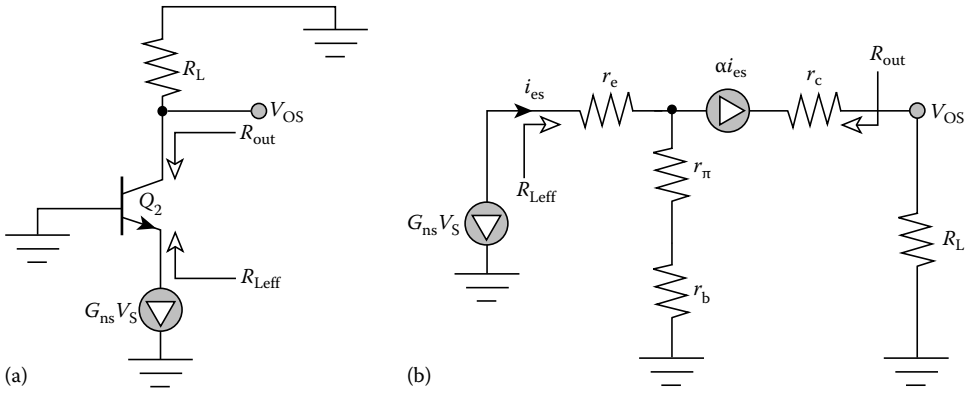


FIGURE 2.30 (a) The effective AC schematic diagram of the common-base component of the common-emitter–common-base cascode of Figure 2.28. (b) Small-signal model of the AC circuit in (a).

approaches $(\beta + 1)r_o$, since the terminating resistance, R_{ns}, seen in the emitter circuit of transistor Q_2 is of the order of r_o.

Equation 2.63 is similar in form to Equation 2.41 which defines the voltage gain of a simple common-emitter amplifier. A careful comparison of the two subject relationships suggests that the voltage gain of the common-emitter–common-base cascode of Figure 2.28 is equivalent to the voltage gain achieved by a simple common-emitter stage, whose output port is loaded in an effective load resistance of αR_L. Although α is close to unity, but nonetheless always less than one, an effective load of αR_L implies that the voltage gain of the cascode is slightly less than that achieved by the common-emitter stage alone, provided, of course, that the transistors utilized in both configurations have identical small-signal parameters. A question therefore arises as to the prudence of incorporating common-base signal processing in conjunction with a common-emitter unit stage.

In fact, no practical purpose is served by a common-emitter–common-base cascode if the load resistance R_L driven by the amplifier is very small. But if the load resistance imposed on the output port of a simple common-emitter amplifier is large, as it is when the load itself is realized actively, as per Example 2.2, the effective transistor beta defined by Equation 2.42 is appreciably smaller than the actual small-signal beta. The result is a degraded common-emitter amplifier voltage gain. In this situation, the insertion of a common-base stage between the output port of the common-emitter amplifier and R_L, as diagrammed in Figure 2.28, increases the degraded gain of the common-emitter amplifier alone by restoring the effective beta of the common-emitter transistor to a value that approximates the actual small-signal beta of the transistor. This observation follows from the fact that the effective load resistance, R_{Leff}, seen by the collector of the common-emitter transistor in the cascode topology is of the order of only a small diode resistance. It follows from Equation 2.42 that β_{eff} for $R_{LT} = R_{Leff}$ is likely to significantly larger than the value of β_{eff} that derives from the load condition, $R_{LT} = R_L$.

The reason for using common-base circuit technology in conjunction with a common-emitter amplifier is circuit broadbanding. In particular, a carefully designed common-emitter–common-base cascode configuration displays a 3 dB bandwidth and a gain-bandwidth product that are significantly larger than the bandwidth and gain-bandwidth product afforded by a common-emitter stage of comparable gain. The primary reason underlying this laudable attribute is the low effective load resistance presented to the collector of the common-emitter stage by the emitter of the common-base structure. This low resistance attenuates the magnitude of the phase-inverted voltage gain of the common-emitter circuit, thereby reducing the deleterious effects of Miller multiplication of the base-collector junction depletion capacitance implicit to the common-emitter transistor [8].

2.2.3.4 Common-Collector Amplifier

While the common-base configuration functions as a current buffer, the ''common-collector amplifier,'' or ''emitter follower,'' with AC schematic diagrams that appear in Figure 2.31a and b operates as a voltage buffer. It offers high input resistance, low output resistance, a voltage gain approaching unity, and a moderately large current gain. The small-signal model of the emitter follower is the circuit in Figure 2.31c.

Assuming infinitely large Early resistance, r_o, a straightforward analysis of the subject model reveals an emitter-follower voltage gain A_{vcc} of

$$A_{vcc} = \frac{V_{OS}}{V_S} = \frac{R_{LT}}{R_{LT} + R_d + (1-\alpha)R_{ST}} \tag{2.64}$$

where R_d is the diode resistance given by Equation 2.33. Observe that A_{vcc} is a positive less than unity number. The indicated gain approaches one for $R_{LT} \gg R_d + (1-\alpha)R_{ST}$. Although the voltage gain is less than one, the corresponding current gain, which is simply the voltage gain scaled by a factor of (R_{ST}/R_{LT}), can be substantially larger than one.

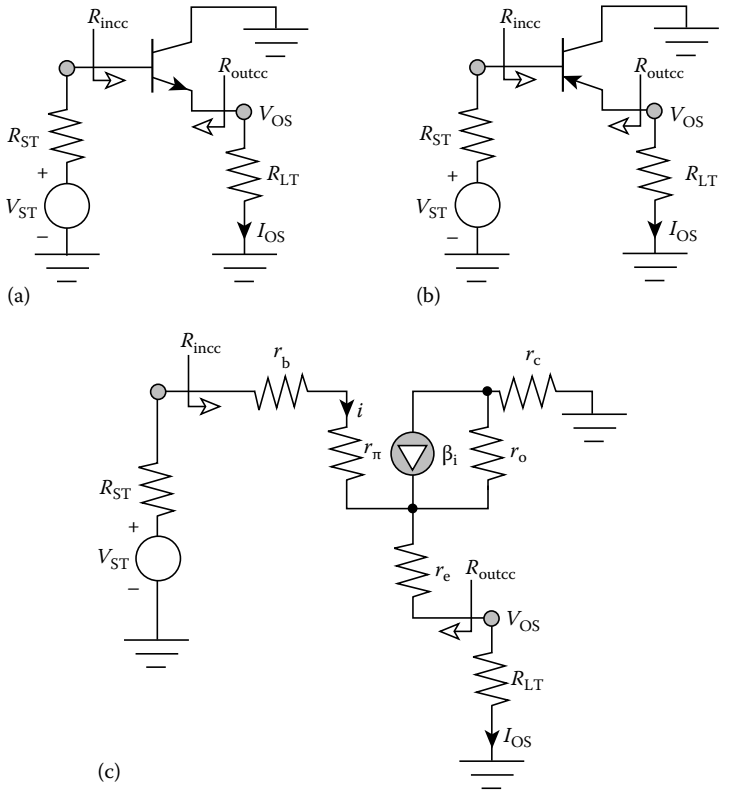


FIGURE 2.31 (a) AC schematic diagram of an NPN common-collector amplifier. (b) AC schematic diagram of PNP common-collector amplifier. (c) Small-signal low-frequency equivalent circuit of the common-collector amplifier, assuming that the Early resistance is sufficiently large to ignore.

It is simple to confirm that the driving point input resistance R_{incc} and the driving point output resistance R_{outcc} of the emitter follower are, respectively,

$$R_{incc} = r_b + r_\pi + (\beta + 1)(r_e + R_{LT}) = (\beta + 1)(R_d + R_{LT}) \tag{2.65}$$

and

$$R_{outcc} = r_e + \frac{r_\pi + r_b + R_{ST}}{(\beta + 1)} = R_d + (1 - \alpha)R_{ST} \tag{2.66}$$

It is interesting and instructive to note that the driving point input resistance, R_{incc}, of an emitter follower is of the same form as the driving point input resistance, R_{incc}, of a common-emitter stage. Indeed, if R_{LT} in the emitter follower is zero, $R_{incc} + R_{ince}$. This result is reasonable in view of the fact that for both the emitter-follower and common-emitter configurations, the input resistance is, as suggested by the test circuits in Figure 2.32a and b, the Thévenin resistance presented to the source circuit by the base of the subject transistor. Moreover, the common-collector output resistance R_{outcc} mirrors the driving point input resistance R_{incb} for the common-base amplifier. In fact, $R_{incb} + R_{outcc}$ if the base of a common-base amplifier is terminated to ground in a resistance of value R_{ST}. Once again, the latter observation is

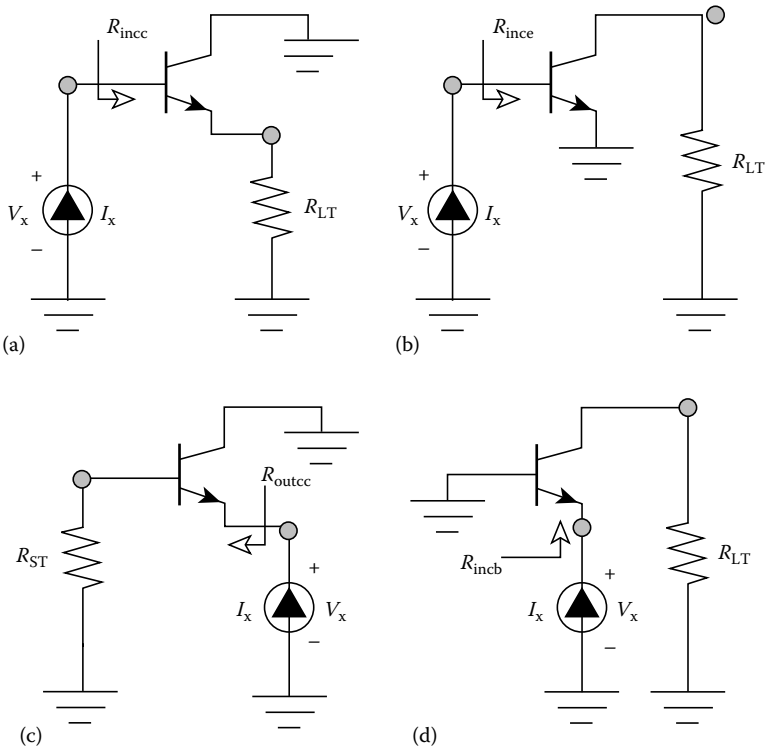


FIGURE 2.32 (a) Test circuit for determining the driving point input resistance R_{incc} of a common-collector amplifier. (b) Test circuit for determining the driving point input resistance R_{ince} of a common-emitter amplifier. Note that $R_{incc} = R_{ince}$ if R_{LT} in the common-collector unit is zero. (c) Test circuit for determining the driving point output resistance R_{outcc} of a common-collector amplifier. (d) Test circuit for determining the driving point input resistance R_{incb} of a common-base amplifier. Note that $R_{outcc} = R_{incb}$ if R_{ST} in the common-collector unit is zero.

intuitively correct, for, as depicted in Figure 2.32c and d, the emitter of the transistor comprises the input terminal for the common-base amplifier and the output terminal of an emitter follower.

An inspection of Equation 2.64 confirms a common-collector voltage gain that tends toward unity for progressively larger Thévenin load resistances, R_{LT}. Correspondingly, the driving point input resistance of an emitter follower increases dramatically with increasing R_{LT}. These observations are often pragmatically exploited by supplanting the passive load in the schematic diagram of Figure 2.31a with an active load, as suggested in Figure 2.33a. Since the effective AC load resistance must be large to achieve a near unity voltage gain, this active load must function as a sink of nominally constant current. To this end, the active load in question is shown conducting a net current that consists of a static current component I_{CS} and a signal current component I_{OS}, where the constant current sink nature of the load implies $I_{CS} \gg I_{OS}$.

The subject active load can be represented by its Norton equivalent circuit, as diagrammed in Figure 2.33b, where I_{CS} is depicted as a constant current source and I_{OS} is made to flow through a resistance, R_{cs}. The latter branch element represents the dynamic resistance presented to the emitter-follower output port by the two terminal active termination. Note that $R_{cs} = \infty$ yields $I_{os} = 0$, which implies an active load that behaves as an ideal constant current sink. The corresponding AC schematic

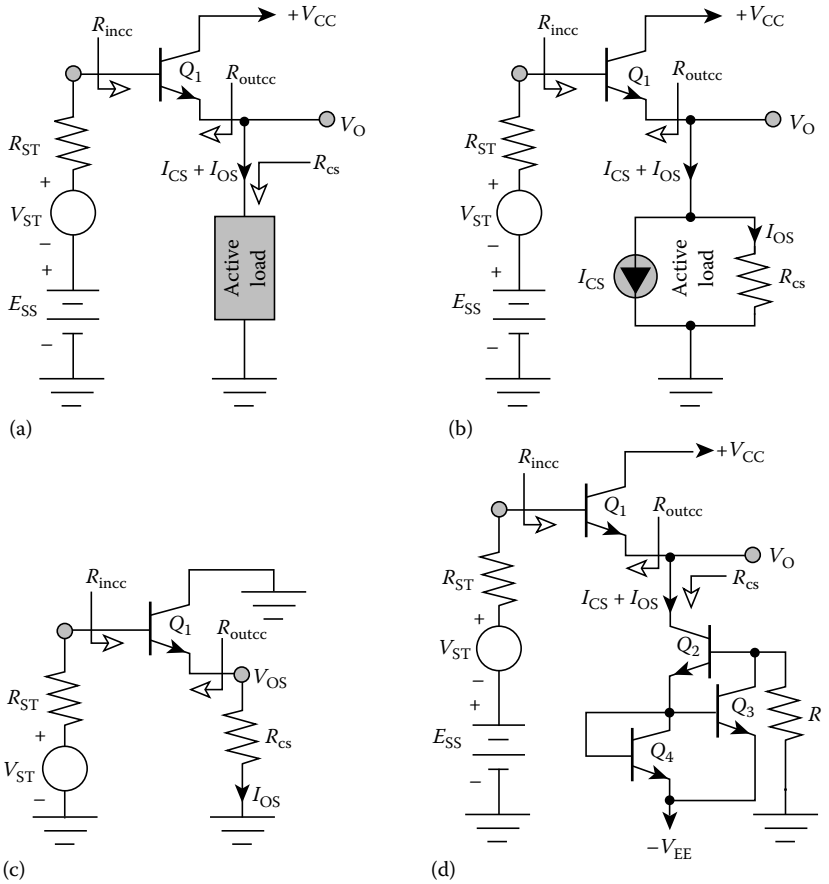


FIGURE 2.33 (a) Emitter follower with active load that conducts a static current I_{CS} and a signal component I_{OS}. (b) The Norton equivalent circuit of the active load. (c) AC schematic diagram of the actively loaded emitter follower. (d) Wilson current mirror realization of the active load.

diagram, which is offered in Figure 2.33c, is derived from Figure 2.33b by setting I_{CS} to zero, since I_{CS} itself is a constant current that is devoid of any signal component. Additionally, the biasing sources, V_{CC}, V_{EE}, and E_{SS} are presumed ideal and are therefore set to zero as well.

The AC schematic diagram in Figure 2.33c is identical to that in Figure 2.31a, subject to the proviso that $R_{LT} = R_{cs}$. But since R_{cs} is presumably a large resistance, substituting $R_{LT} = R_{cs}$ into Equation 2.66 to evaluate the voltage gain of the actively loaded emitter follower is at least theoretically inappropriate because the subject gain equation is premised on the assumption of a large Early resistance, r_o; specifically Equation 2.64 reflects the assumption $R_{LT} \ll r_o$. A voltage gain expression, more accurate than Equation 2.64, derives from an analysis of the model in Figure 2.31c. If r_o is included in this analysis, but if r_e and r_c are sufficiently small to justify their neglect, the result for $R_{LT} = R_{cs}$ is

$$A_{vcc} = \frac{V_{OS}}{V_S} \approx \frac{r_o \| R_{cs}}{(r_o \| R_{cs}) + R_d + (1 - \alpha) R_{ST}} \tag{2.67}$$

where the algebraic form collapses to Equation 2.64 if $R_{cs} \ll r_o$. Similarly, the revised expression for the driving point input resistance is

$$R_{incc} \approx r_b + r_\pi + (\beta + 1)(r_o \| R_{cs}) \tag{2.68}$$

The output resistance R_{outcc} remains as stipulated by Equation 2.66.

The active load appearing in Figure 2.33a can be realized as any one of a variety of NPN current sources [9]. Figure 2.33d offers an examples of such a realization in the form of the Wilson current mirror formed of transistors Q_2, Q_3, and Q_4, and the current setting resistor, R [10]. This subcircuit establishes an extremely high dynamic resistance between the collector of transistor Q_2 and the signal ground. In particular, if β_2 and r_{o2} symbolize the AC beta and forward Early resistance, respectively, of transistor Q_2, it can be shown that

$$R_{cs} \approx \frac{\beta_2 r_{o2.}}{2} \tag{2.69}$$

Note further that the static current, I_{CS}, conducted by the Wilson mirror flows through the emitter lead of the emitter-follower transistor Q_1. Thus, the biasing stability of Q_1 is determined by the thermal sensitivity of the static current that flows in the Wilson subcircuit.

Example 2.3

In order to dramatize the voltage buffering property of an emitter follower, return to the amplifier addressed analytically in Example 2.1 and drawn schematically in Figure 2.22a. Let the two coupling capacitors remain large enough to approximate them as AC short circuits over the signal frequency range of interest, and let the small-signal parameters of the two transistors remain at $r_b = 90\ \Omega$, $r_c = 55\ \Omega$, $r_e = 1.5\ \Omega$, $r_\pi = 970\ \Omega$, $r_o = 42\ k\Omega$, and $\beta = 115$. The circuit parameters also remain the same; namely, $R_1 = 2.2\ k\Omega$, $R_2 = 1.3\ k\Omega$, $R_{EE} = 75\ \Omega$, $R_{CC} = 3.9\ k\Omega$, and $R_S = 300\ \Omega$. But instead of $R_L = 1.0\ k\Omega$, consider an external load termination of $R_L = 300$. Reevaluate the small-signal voltage gain, $A_v = V_{OS}/V_S$, for the subject amplifier. Compare this result to the voltage gain achieved when an emitter follower is inserted between the collector of transistor Q_1 and the 300 Ω load termination, as depicted in Figure 2.34a. Assume that the small-signal parameters of the emitter-follower transistor Q_3 and those of the two transistors that comprise the diode-compensated current sink load of the follower are identical to the model parameters of transistors Q_1 and Q_2.

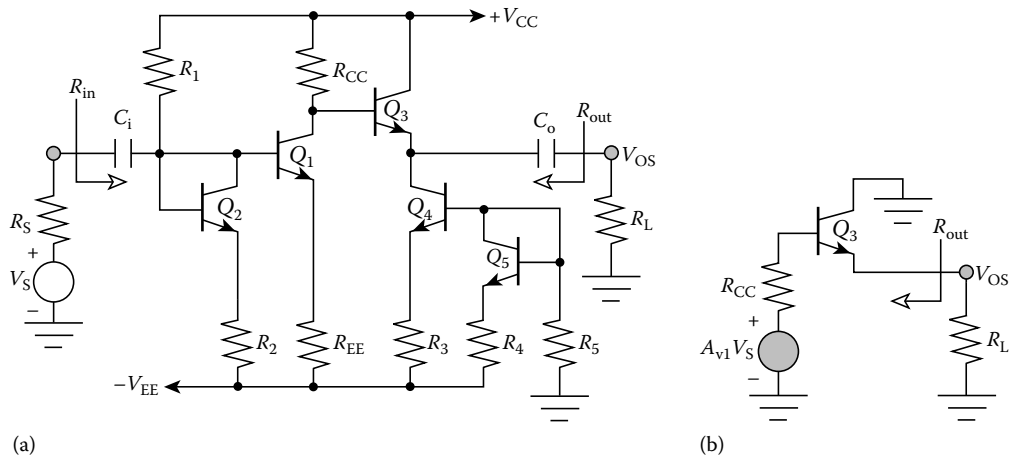


FIGURE 2.34 (a) The amplifier of Figure 2.22, but with an emitter-follower buffer inserted between the gain stage and the terminating load resistance. (b) AC schematic diagram of the amplifier in (a).

Solution

1. With reference to Figure 2.22 and Example 2.1, the Thévenin load resistance R_{LT} is now

$$R_{LT} = R_{CC} \| R_L = 278.6\ \Omega$$

The parameters R_{ST} and K_{ST} in Figure 2.22c remain unchanged at the previously computed values of $R_{ST} = 219.7\ \Omega$ and $K_{ST} = 0.733$. Then, ignoring the effects of the finite, but large, Early resistance r_o, the voltage gain is

$$A_v = \frac{V_{OS}}{V_S} \approx -\frac{\beta K_{ST} R_{LT}}{R_{ST} + r_b + r_\pi + (\beta + 1)(r_e + R_{EE})}$$

whence $A_v = -2.31$ V/V.

2. Consider now the amplifier modification shown in Figure 2.34a. Transistor Q_3 functions as an emitter follower to buffer the terminating load resistance R_L effectively seen by the gain stage formed of transistor Q_1 and its peripheral elements. Transistors Q_3 and Q_4 form a diode-compensated current sink that comprises the active load presented to the emitter-follower output port under static operational conditions. To the extent that r_o can be tacitly ignored, this current sink comprises an infinitely large dynamic resistance. Accordingly, the AC schematic diagram seen to the right of the collector of transistor Q_1 is the structure identified in Figure 2.34b.
3. The source circuit that drives the base of transistor Q_3 in Figure 2.34b is the Thévenin equivalent circuit established at the collector of transistor Q_1 in Figure 2.34a. The signal voltage associated with this source circuit is the open circuit voltage developed at the Q_1 collector; that is, it is the voltage at the Q_1 collector with the load formed of transistor Q_3 and its peripheral elements removed. Since the circuit to the left of the base of transistor Q_3 is a linear network, this Thévenin voltage is necessarily proportional to the input signal V_S. The indicated constant of proportionality in Figure 2.34b, A_{v1}, can rightfully be termed the open-circuit voltage gain of the first stage of the subject amplifier. This is to say that A_{v1} is A_v, as determined in step 1 above, but with R_L removed and therefore, R_{LT} set equal to R_{CC}. It follows that

$$A_{v1} = -\frac{\beta K_{ST} R_{CC}}{R_{ST} + r_b + r_\pi + (\beta + 1)(r_e + R_{EE})}$$

or $A_{v1} = -32.38$ V/V. Since r_o is taken to be infinitely large, the resistance seen looking into the collector of transistor Q_1 is likewise infinitely large. As a result, the Thévenin resistance associated with the source circuit in the AC diagram of Figure 2.34b is $R_{CC} = 3.9$ kΩ.

4. Recalling Equation 2.64, the voltage gain of the circuit in Figure 2.22b is

$$A_{vcc} = \frac{V_{OS}}{A_{v1} V_S} = \frac{R_L}{R_L + R_d + (1 - \alpha) R_{CC}} = 0.871 \text{ V/V}$$

The resultant overall circuit gain is

$$A_v = \frac{V_{OS}}{V_S} = A_{v1} A_{vcc}$$

or $A_v = -28.21$ V/V. Recalling the results of step 1 of this computational procedure, the effect of the emitter follower is to boost the gain magnitude of the original configuration by a factor of about 12.2.

5. From Equation 2.66, the driving point output resistance of the buffered amplifier is

$$R_{out} = R_d + (1 - \alpha) R_{CC}$$

or $R_{out} = 44.3$ Ω. Note that for the original nonbuffered case, the output resistance is $R_{CC} = 3.9$ kΩ.

2.2.3.5 Darlington Connection

In the Darlington connection, whose basic schematic diagram is abstracted in Figure 2.35a, the emitter of one transistor Q_1 is incident with the base of a second transistor Q_2 and the two transistor collector leads are connected. The output signal is extracted as either the current flowing in the collector of transistor Q_2 or the voltage developed at the emitter of Q_2. In the former case, the indicated Darlington connection functions as a transconductance amplifier. In the latter case, an output signal voltage at the emitter of Q_2 renders the connection functional as a voltage follower, or buffer. In both applications, the small-signal driving point input resistance, R_{ind}, seen looking into the base of transistor Q_1 is large. On the other hand, the driving point output resistance, R_{outed}, seen at the emitter is small and virtually independent of the source resistance R_S. The output resistance, R_{outcd}, presented to the node at which the two transistor collectors are incident is large. At the expense of forward transconductance, R_{outcd} can be enhanced by returning the collector of Q_1 to the $+V_{CC}$ bus, instead of to the collector of transistor Q_2. For a nonzero collector load resistance, R_{LC}, this alternate connection, which is diagrammed in Figure 2.35b eliminates Miller multiplication of the base–collector junction capacitance of transistor Q_1, thereby resulting in an improved transconductance frequency response.

A fundamental problem that plagues both of the foregoing Darlington connections is the fact that the static emitter current conducted by Q_1 is identical to the static base current drawn by Q_2. Accordingly, the emitter current of Q_1 is likely to be much smaller than the biasing current commensurate with optimal gain-bandwidth product in this device. Moreover, this emitter current cannot be predicted accurately since it is inversely proportional to the Q_2 static beta, whose numerical value is an unavoidable uncertainty. This poor biasing translates into an unreliable delineation of the static and small-signal parameters for Q_1. In turn, potentially significant uncertainties shroud the forward transfer and driving point resistance characteristics of the Darlington configuration.

To remedy the situation at hand, an additional current path, usually directed to signal ground, is provided at the junction of the Q_1 emitter and the Q_2 base, as suggested in Figure 2.35c and d.

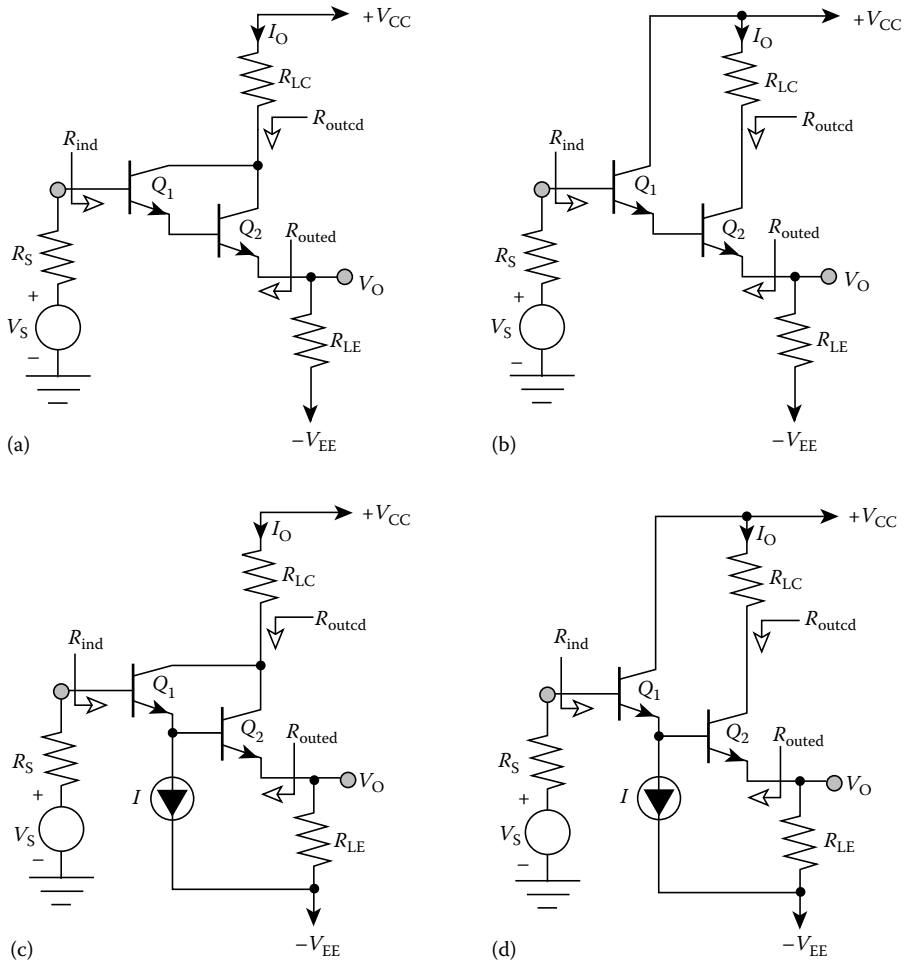


FIGURE 2.35 (a) The basic Darlington connection. (b) Alternative Darlington connection for wide-band transconductance response. (c) Darlington connection with input transistor current compensation. (d) Alternative Darlington connection with input transistor current compensation.

The appended current path can be a simple two terminal resistance, although care must be exercised to ensure that this resistance is sufficiently large to avoid seriously compromising the large driving point input resistance afforded by the basic Darlington connection in either of the preceding diagrams. Since large resistance and realistic biasing currents may prove to be conflicting design requirements, the appended current path is often an active current sink, such as the Wilson mirror load explored earlier in conjunction with the common-collector amplifier. Note in the latter two diagrams that the current, indicated as I, conducted by the appended passive or active current path is essentially the emitter current of transistor Q_1, provided that I is much larger than the base currents of Q_2.

The small-signal BJT equivalent circuit of Figure 2.16a can be used to deduce the transfer and driving point resistance characteristics of any of the Darlington connections depicted in Figure 2.35. An analysis is provided herewith for only the configuration in Figure 2.35c, since this topology is the most commonly encountered Darlington circuit and the others are amenable to very straightforward analyses. To this end, the model for the subject structure is offered in Figure 2.36, where it is assumed that both transistors are biased so that their corresponding small-signal parameters are nominally identical. Moreover, the Early

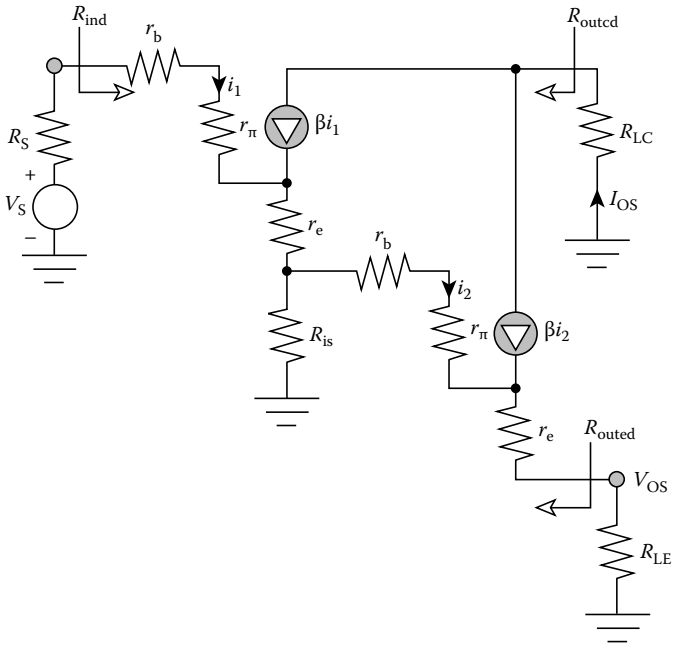


FIGURE 2.36 Small-signal equivalent circuit of the Darlington connection in Figure 2.35c. The Early resistance is ignored, and both transistors are presumed to have identical corresponding small-signal parameters. The resistance, R_{is}, represents the terminal AC resistance associated with the appended current path conducting current I in Figure 2.35c.

resistance of each transistor is presumed to be sufficiently large to warrant its neglect, and a resistance, R_{is}, is included to account for the terminal resistance of the appended current path discussed above. Letting

$$k_{is} \triangleq \frac{R_{is}}{R_{is} + (\beta + 1)(R_d + R_{LE})} \tag{2.70}$$

denote the small-signal current divider between the appended current path and the base circuit of transistor Q_2, it can be shown that the driving point input resistance R_{ind} is

$$R_{ind} = (\beta + 1)[R_d + (\beta + 1)k_{is}(R_d + R_{LE})] \tag{2.71}$$

For large ac beta,

$$R_{ind} \approx (\beta + 1)^2 k_{is}(R_d + R_{LE}) \tag{2.72}$$

which is maximal for $k_{is} \approx 1$. From Equation 2.70, the latter constraint mandates that the appended current path be designed so that its small-signal terminal resistance satisfies the inequality $R_{is} \gg (\beta + 1)(R_d + R_{LE})$.

The voltage gain, A_{vd}, from the signal source to the emitter port is

$$A_{vd} = \frac{V_{OS}}{V_S} = \frac{(\beta + 1)^2 k_{is} R_{LE}}{R_S + (\beta + 1)R_d + (\beta + 1)^2 k_{is}(R_d + R_{LE})} \tag{2.73}$$

where V_{OS} is the signal component of the net output voltage V_O. Equation 2.53 reduces to

$$A_{vd} \approx \frac{R_{LE}}{R_d + R_{LE}} \tag{2.74}$$

for large ac beta. The corresponding driving point output resistance, R_{outed}, is

$$R_{outed} = R_d + \frac{R_d}{(\beta + 1)k_{is}} + \frac{R_S}{(\beta + 1)^2 k_{is}} \approx R_d \tag{2.75}$$

At the collector port, the driving point output resistance, R_{outcd}, is infinitely large to the extent that the Early resistance r_o of both transistors can be ignored. For finite r_o, this resistance is of the order of, and slightly larger than, $(r_o/2)$. Finally, the model in Figure 2.36 yields a forward transconductance, G_{fd}, from the signal source to the collector port of

$$G_{fd} = \frac{I_{OS}}{V_S} = \frac{\beta[1 + (\beta + 1)k_{is}]}{R_S + (\beta + 1)R_d + (\beta + 1)^2 k_{is}(R_d + R_{LE})} \tag{2.76}$$

where I_{OS} is the signal component of the net output current I_O. Equation 2.76 collapses to

$$G_{fd} \approx \frac{\alpha}{R_d + R_{LE}} \tag{2.77}$$

for large AC beta.

2.2.4 Differential Amplifier

The ''differential amplifier'' is a four-port network, as suggested in Figure 2.37a. Source signals represented by the voltages, V_{S1} and V_{S2}, which have Thévenin resistances of R_{S1} and R_{S2}, respectively, are applied to the two amplifier input ports. The two output ports are terminated in three load resistances. Two of these loads, R_{L1} and R_{L2}, are ''single-ended terminations'' in that they provide a signal path to ground from each of the two output terminals. A third load resistance, R_{LL}, is differentially connected between the two output terminals. In response to the two applied source signals, two ''single-ended output'' voltages, V_{O1} and V_{O2}, are generated across R_{L1} and R_{L2}, and a ''differential output voltage,'' V_{DO}, is established across R_{LL}. This third output response is the difference between V_{O1} and V_{O2}; that is,

$$V_{DO} = V_{O1} - V_{O2} \tag{2.78}$$

The salient features of a differential amplifier are unmasked by the concepts of ''differential- and common-mode'' excitation and response. To this end, let the ''differential input source voltage,'' V_{DI}, be defined as

$$V_{DI} \triangleq V_{S1} - V_{S2} \tag{2.79}$$

and let the ''common-mode input voltage,'' V_{CI}, be

$$V_{CI} \triangleq \frac{1}{2}(V_{S1} + V_{S2}) \tag{2.80}$$

The differential input voltage is seen as the difference between the two applied source excitations. On the other hand, the common-mode input voltage is the arithmetic average of the two source voltages.

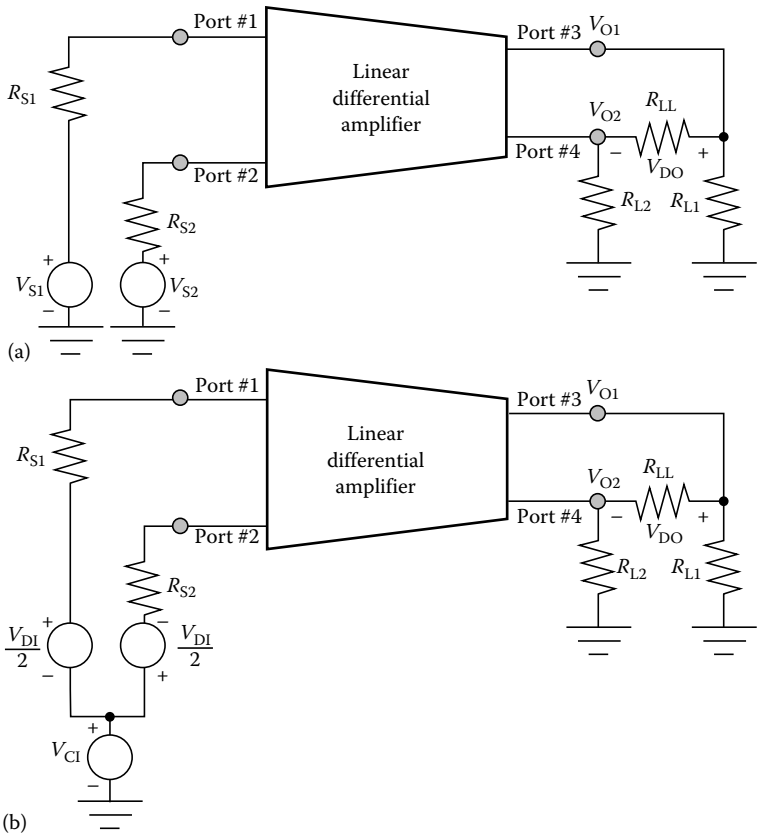


FIGURE 2.37 (a) System-level diagram of a differential amplifier. (b) System-level diagram that depicts the electrical implications of the common-mode and the differential-mode input source voltages.

When solved for V_{S1} and V_{S2}, Equations 2.79 and 2.80 give

$$V_{S1} = V_{CI} + \frac{1}{2}V_{DI} \tag{2.81a}$$

$$V_{S2} = V_{CI} - \frac{1}{2}V_{DI} \tag{2.81b}$$

The preceding two expressions allow the diagram of Figure 2.37a to be drawn in the form shown in Figure 2.37b. This alternative representation underscores the fact that the Thévenin voltage applied to either input port is the superposition of a common-mode source voltage and a component proportional to the differential-mode source voltage. The common-mode component raises both of the open-circuit input terminals to a voltage that lies above ground by an amount, V_{CI}. Superimposed with V_{CI} at the open circuit terminals of port 1 is a differential-mode voltage, $V_{DI}/2$. Simultaneously, a voltage of $-V_{DI}/2$ superimposes with V_{CI} at the open-circuit terminals of port 2.

The fact that two general source excitations applied to a four-port system can be separated into a voltage component that appears only differentially across the two system input ports and a single-ended common-mode voltage component that is simultaneously incident with both of the system input ports makes it possible to achieve signal discrimination in a differential circuit. In particular, a differential amplifier can be designed so that it amplifies the differential component of two source signals while

rejecting (in the sense of amplifying with near zero gain) their common-mode component. Signal discrimination is useful whenever an electronic system must process low-level electrical signals that are contaminated by spurious inputs, such as the voltage ramifications of electromagnetic interference or the biasing perturbations induced by temperature. If the two input ports of a differential amplifier are geometrically proximate and have matched driving point input impedances, these spurious excitations impact the two input ports identically. The undesired inputs are therefore common-mode excitations that can be rejected by a differential amplifier that is well designed in the sense of producing output port responses that are sensitive to only differential inputs.

If the differential amplifier in Figure 2.37 is linear, superposition theory gives

$$V_{O1} = A_{11}V_{S1} + A_{12}V_{S2} \tag{2.82a}$$

$$V_{O2} = A_{21}V_{S1} + A_{22}V_{S2} \tag{2.82b}$$

where the A_{ij} are constants, independent of V_{S1} and V_{S2}. When Equation 2.81a and b are inserted into the last two relationships, the single-ended output voltages are expressible as

$$V_{O1} = (A_{11} + A_{12})V_{CI} + (A_{11} - A_{12})\frac{V_{DI}}{2} \tag{2.83a}$$

$$V_{O2} = (A_{22} + A_{21})V_{CI} - (A_{22} - A_{21})\frac{V_{DI}}{2} \tag{2.83b}$$

It follows that the differential output voltage is

$$V_{DO} = (A_{11} - A_{22} + A_{12} - A_{21})V_{CI} + (A_{11} + A_{22} - A_{12} - A_{21})\frac{V_{DI}}{2} \tag{2.84}$$

Since the common-mode output voltage is

$$V_{CO} \triangleq \frac{1}{2}(V_{O1} + V_{O2}) \tag{2.85}$$

Equation 2.83a and b yield

$$V_{CO} = \left(\frac{A_{11} + A_{22} + A_{12} + A_{21}}{2}\right)V_{CI} + \left(\frac{A_{11} - A_{22} - A_{12} + A_{21}}{2}\right)\frac{V_{DI}}{2} \tag{2.86}$$

The ability of a differential amplifier to process differential excitations is measured by the ''differential-mode voltage gain,'' A_D. This performance index is defined as the ratio of the differential output voltage to the differential input voltage, under the condition of zero common-mode input voltage. From Equation 2.84,

$$A_D \triangleq \left[\frac{V_{DO}}{V_{DI}}\right]\Bigg|_{V_{CI}=0} = \frac{A_{11} + A_{22} - A_{12} - A_{21}}{2} \tag{2.87}$$

On the other hand, the ''common-mode voltage gain,'' A_C, is a measure of the common-mode signal rejection characteristics of a differential amplifier. It is the ratio of the common-mode output voltage to the common-mode input voltage, under the condition of zero differential input voltage. Using Equation 2.86,

$$A_C \triangleq \left[\frac{V_{CO}}{V_{CI}}\right]\Bigg|_{V_{DI}=0} = \frac{A_{11} + A_{22} + A_{12} + A_{21}}{2} \tag{2.88}$$

A measure of the degree to which a differential amplifier rejects common-mode excitation is the "common-mode rejection ratio ρ," which is the ratio of the differential-mode voltage gain to the common-mode voltage gain. From Equations 2.87 and 2.88

$$\rho \triangleq \frac{A_D}{A_C} = \frac{A_{11} + A_{22} - A_{12} - A_{21}}{A_{11} + A_{22} + A_{12} + A_{21}} \tag{2.89}$$

A common-mode gain of zero indicates that no common-mode output results from the application of common-mode input signals. Therefore, a practical design goal is the realization of a differential amplifier that has the largest possible magnitude of common-mode rejection ratio.

2.2.4.1 Balanced Differential Amplifier

Most differential amplifiers are "balanced." Two operating requirements are satisfied by balanced differential systems. First, with zero common-mode input voltage, the two single-ended output voltages are mutually phase inverted, but otherwise identical. By Equation 2.83a and b, the balance requirement implies the parametric constraint.

$$A_{11} - A_{12} = A_{22} - A_{21} \tag{2.90}$$

Second, equal single-ended output voltages result when the differential-mode input voltage is zero. Using Equation 2.83a and b once again, this stipulation requires

$$A_{11} + A_{12} = A_{22} + A_{21} \tag{2.91}$$

Equations 2.90 and 2.91 combine to deliver the balanced operating requirement

$$\left.\begin{aligned} A_{11} &= A_{22} \\ A_{12} &= A_{21} \end{aligned}\right\} \tag{2.92}$$

From Equations 2.87 through 2.89, the differential-mode voltage gain, the common-mode voltage gain, and the common-mode rejection ratios of a balanced differential amplifier are

$$A_D = A_{11} - A_{12} \tag{2.93}$$

$$A_C = A_{11} + A_{12} \tag{2.94}$$

$$\rho = \frac{A_{11} - A_{12}}{A_{11} + A_{12}} \tag{2.95}$$

Moreover, the single-ended output voltages in Equation 2.83a and b become

$$V_{O1} = A_C V_{CI} + A_D \frac{V_{DI}}{2} = \frac{A_D}{2}\left(1 + \frac{2V_{CI}}{\rho V_{DI}}\right) V_{DI} \tag{2.96a}$$

$$V_{O2} = A_C V_{CI} - A_D \frac{V_{DI}}{2} = -\frac{A_D}{2}\left(1 - \frac{2V_{CI}}{\rho V_{DI}}\right) V_{DI} \tag{2.96b}$$

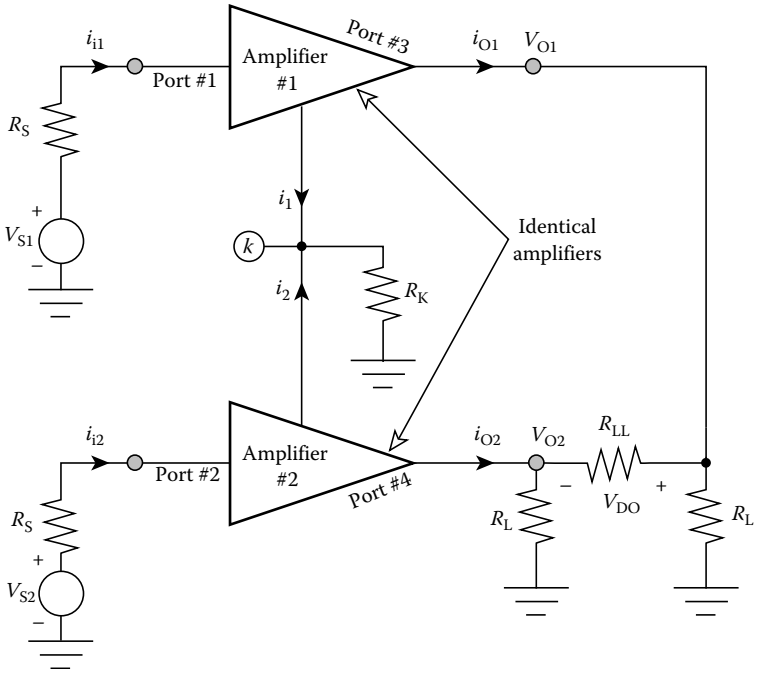


FIGURE 2.38 Generalized system diagram of a balanced differential system. The topology is an AC schematic diagram in that requisite biasing subcircuits of either amplifier are not shown.

which give rise to a differential response of

$$V_{DO} = V_{O1} - V_{O2} = A_D V_{DI} \tag{2.97}$$

Equation 2.96a and b shows that a balanced differential amplifier having a very large common-mode rejection ratio produces single-ended outputs that are nominally phase-inverted versions of one another and approximately independent of the common-mode input voltage. On the other hand, the differential output voltage of a balanced system is independent of the common-mode input signal, regardless of the value of the common-mode rejection ratio.

Figure 2.38 depicts the most straightforward way to implement balance in a differential configuration. In this abstraction, two identical single-ended amplifiers, such as those discussed in earlier subsections, are interconnected to establish signal flow paths between single-ended input and single-ended output ports. This topology boasts integrated circuit practicality, since it exploits the inherent ability of a mature monolithic fabrication process to produce well-matched equivalent components. The two single-ended amplifiers in the subject figure are topologically identical, and they incorporate matched active devices that are biased at the same quiescent-operating points. Thus, amplifiers 1 and 2 have small-signal two-port equivalent circuits that are reflective of one another. In order to ensure balanced operation, the single-ended output ports of each amplifier are terminated to ground in equal load resistances R_L. Similarly, the Thévenin source resistances are equivalent. Observe that balance implies that the upper half of the circuit in Figure 2.38 is a mirror image of the lower half of the system schematic diagram. This interpretation begets the common reference to a balanced differential amplifier as a ''differential pair.''

The balance condition entails the following engineering constraints:

1. Under the case of differential-mode excitation, which implies $V_{S1} = -V_{S2} = V_{DI}/2$ and hence, $V_{CI} = 0$, the currents indicated in Figure 2.38 are such that $i_{i1} = -i_{i2}$, $i_{01} = -i_{02}$, and $i_1 = -i_2$.

Since the resistance, R_K, conducts a current equal to the sum of i_1 and i_2, $i_1 = -i_2$ clamps node k to signal ground potential for exclusively differential-mode inputs. Moreover, Equation 2.96a and b confirm $V_{01} = -V_{02} = A_D\ V_{DI}/2$, which produces a signal current through the differential load resistance R_{LL} of

$$\frac{V_{O1} - V_{O2}}{R_{LL}} = \frac{V_{O1}}{R_{LL}/2}$$

Since the single-ended response voltage, V_{O1}, is referred to signal ground, the midpoint of the differential load resistance is effectively grounded for differential inputs.

The foregoing disclosures imply the circuit diagram of Figure 2.39a, which is the so-called "differential-mode half-circuit equivalent" [11,12] of the differential amplifier. For a balanced differential pair driven by exclusively differential inputs, the branch currents, branch voltages, and node voltages computed from an analysis of the structure in Figure 2.39a are precisely the negative of the corresponding circuit variables in the remaining half of the system.

2. Under the case of common-mode inputs, which implies $V_{SI} = V_{S2} = V_{CI}$ and hence, $V_{DI} = 0$, the currents delineated in Figure 2.38 satisfy the constraints, $i_{i1} = i_{i2}$, $i_{01} = i_{02}$, and $i_1 = i_2$. Since the resistance, R_K, conducts a current equal to the sum of i_1 and i_2, $i_1 = i_2$ establishes a voltage at node k of

$$R_K(i_1 + i_2) = (2R_K)i_1$$

that is, the signal voltage developed at node k corresponds to an amplifier 1 current of i_1 flowing through a resistance whose value is twice R_K. Additionally, $V_{O1} = V_{O2} = V_{CO}$, which means that no signal current flows through R_{LL} under exclusively common-mode excitation.

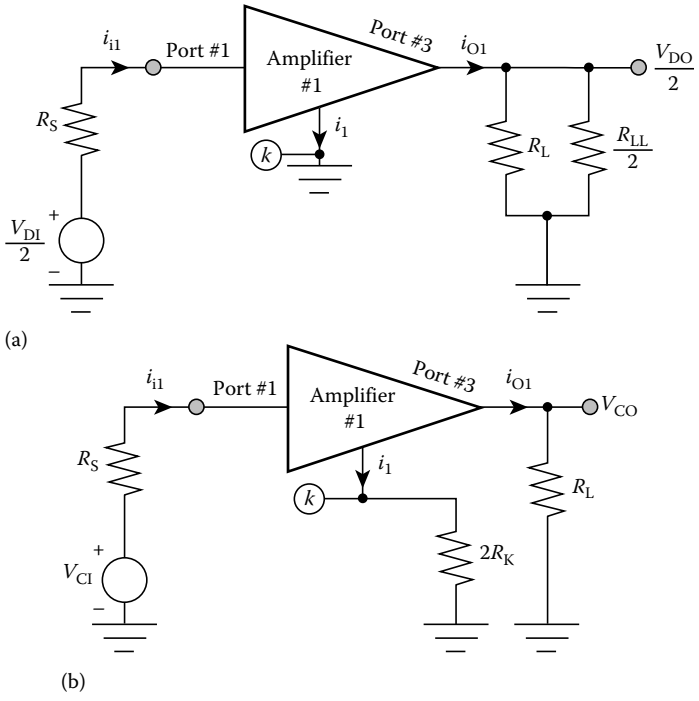


FIGURE 2.39 (a) Differential-mode half-circuit equivalent of the balanced differential amplifier. (b) Common-mode half-circuit equivalent of the balanced differential amplifier.

It follows that the "common-mode half-circuit equivalent" of the differential amplifier is as drawn in Figure 2.39b. For a balanced differential pair driven by exclusively common-mode input signals, the branch currents, branch voltages, and nodal voltages computed for the structure in Figure 2.39b are identical to the corresponding circuit variables in the other half of the circuit.

2.2.4.2 Thévenin Equivalent I/O Circuits

Because the two input ports of a differential amplifier electrically interact with one another, the Thévenin equivalent circuit seen by the two signal sources in Figure 2.37a is itself a two-port network. The branch elements of the Thévenin model are defined in terms of a "differential-mode input resistance," R_{DI}, and a "common-mode input resistance," R_{CI}.

Consider the test circuit of Figure 2.40a, which is configured to formulate the Thévenin equivalent input circuit. This structure is analogous to that of Figure 2.37a, except that the linear differential unit is presumed balanced. Furthermore, the original source circuits are supplanted by the test voltages, V_{tl} and V_{t2}, which establish the input port currents I_{t1} and I_{t2}. Because no signals other than V_{tl} and V_{t2} are applied to the differential pair, the Thévenin equivalent circuit seen between ports 1 and 2 is a resistance, say R_{XI}. Similarly, a second resistance, R_{XX}, is introduced to terminate port 1 to ground. System balance implies that a resistance of the same value terminates the second input port. The hypothesized Thévenin equivalent input circuit is given in Figure 2.40b.

The application of Kirchhoff's current and voltage laws to the model is Figure 2.40b produces

$$I_{t1} - \frac{V_{t1}}{R_{XX}} = \frac{V_{t2}}{R_{XX}} - I_{t2} \tag{2.98a}$$

$$V_{t1} - V_{t2} = R_{XI}\left(I_{t1} - \frac{V_{t1}}{R_{XX}}\right) \tag{2.98b}$$

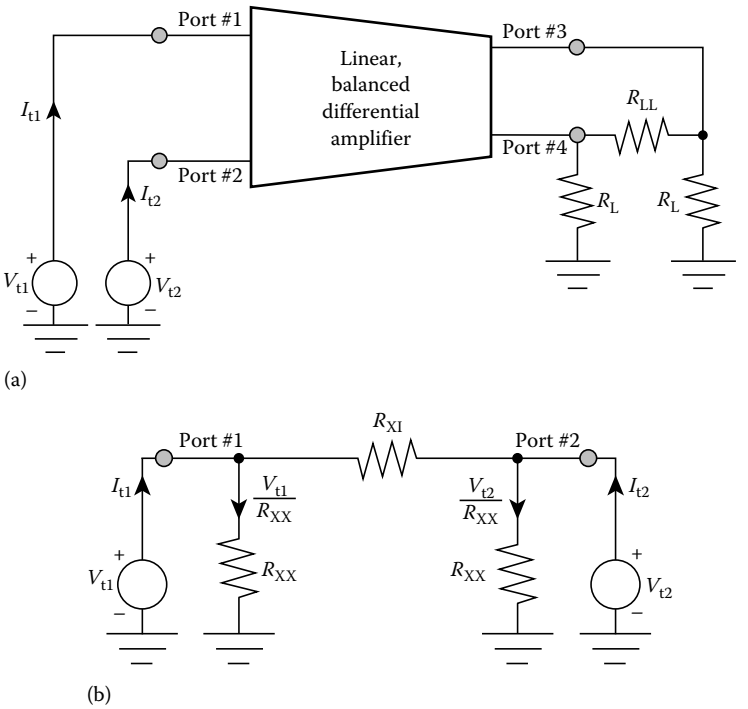


FIGURE 2.40 (a) Test circuit used to evaluate the Thévenin input equivalent circuit of a balanced differential amplifier. Note the connection of balanced loads at ports 3 and 4. (b) Hypothesized Thévenin equivalent input circuit.

If the test voltages, V_{t1} and V_{t2}, are decomposed into their differential (V_{Dt}) and common mode (V_{Ct}) components in accordance with

$$V_{t1} = V_{Ct} + \frac{V_{Dt}}{2} \tag{2.99a}$$

$$V_{t2} = V_{Ct} - \frac{V_{Dt}}{2} \tag{2.99b}$$

Equation 2.98a and b lead to

$$I_{t1} = \frac{V_{Ct}}{R_{XX}} + \frac{V_{Dt}}{R_{XI} \| (2R_{XX})} \tag{2.100a}$$

$$I_{t2} = \frac{V_{Ct}}{R_{XX}} - \frac{V_{Dt}}{R_{XI} \| (2R_{XX})} \tag{2.100b}$$

Equation 2.100a and b implicitly define the common-mode and differential-mode components of the currents I_{t1} and I_{t2} resulting from the test voltages V_{t1} and V_{t2}. Accordingly, the common-mode driving point input resistance, R_{CI}, is

$$R_{CI} = R_{XX} \tag{2.101}$$

On the other hand, the differential-mode driving point input resistance, R_{DI}, is

$$R_{DI} = R_{XI} \| (2R_{CI}) \tag{2.102}$$

where use has been made of Equations 2.100a, b, and 2.101. Note that the model resistance, R_{XI}, is related to the differential input resistance, R_{DI}, of the amplifier by

$$R_{XI} = \frac{2R_{CI}R_{DI}}{2R_{CI} - R_{DI}} \tag{2.103}$$

The test circuits for measuring the common-mode and the differential-mode driving point input resistances derive directly from Equation 2.100a and b. For a differential input test voltage, V_{Dt} of zero, V_{t1} and V_{t2} are identical to the common-mode input test voltage V_{Ct}, whence

$$I_{t1} = \frac{V_{t1}}{R_{XX}} = \frac{V_{t1}}{R_{CI}} \equiv I_{t2} \tag{2.104}$$

It follows that the circuit of Figure 2.41a is appropriate to the measurement (or calculation) of R_{CI} as the Ohm's law ratio of the common-mode test voltage to the resultant common-mode test current. Observe that half circuit analysis measures apply, wherein R_{CI} is the resistance seen looking into port 1, with port 3 terminated to ground in the resistance R_L. For differential testing, $V_{t1} = -V_{t2} = V_{Dt}/2$, whence

$$I_{t1} = \frac{2V_{t1}}{R_{XI} \| (2R_{CI})} = \frac{2V_{t1}}{R_{DI}} = -I_{t2} \tag{2.105}$$

The pertinent test cell is shown in Figure 2.41b. For half-circuit analysis, care should be exercised to recognize that the ratio of the test voltage, V_{t1}, to the corresponding test current, I_{t1}, is one-half of the driving point differential input resistance, R_{DI}. In addition, the subcircuit connecting port 2 to port 4

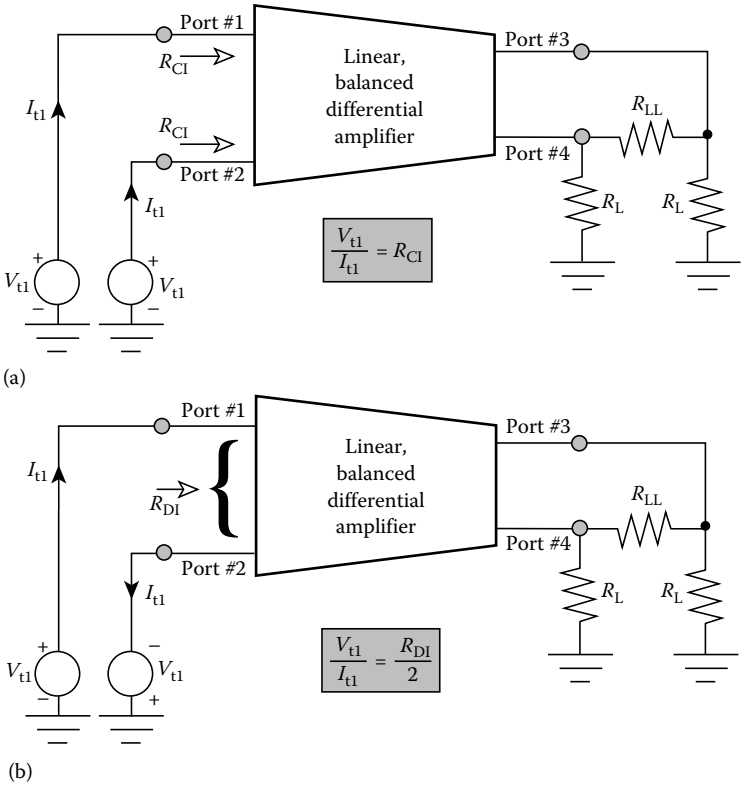


FIGURE 2.41 (a) Test circuit used to evaluate the driving point common-mode input resistance of a linear, balanced differential pair. (b) Test circuit used to evaluate the driving point differential-mode input resistance of a linear, balanced differential pair.

must be removed, and port 3 must be terminated to ground by the shunt interconnection of the resistances, R_L and $R_{LL}/2$.

Just as a Thévenin model can be constructed for the input ports of a balanced differential pair, a Thévenin equivalent circuit can be developed for the output ports. Under zero input conditions, this output model, which is presented in Figure 2.42, is topologically identical to the equivalent circuit in Figure 2.40b. In a fashion that reflects the computation of the resistance parameters for the input equivalent circuit, Figure 2.43a is the test circuit for evaluating the driving point common-mode output resistance, R_{CO}. Figure 2.43b is the test structure for calculating the driving point differential-mode output resistance, R_{DO}. Following Equation 2.103, R_{XO} in Figure 2.42 is given by

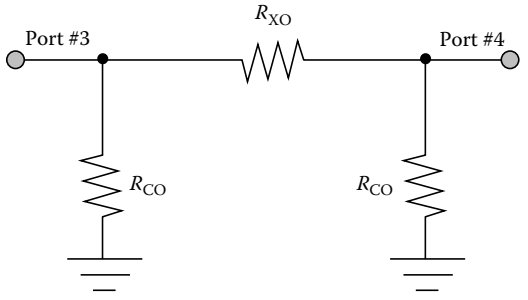


FIGURE 2.42 Thévenin equivalent output circuit of the balanced differential amplifier for the case of zero input signal excitation.

$$R_{XO} = \frac{2R_{CO}R_{DO}}{2R_{CO} - R_{DO}} \tag{2.106}$$

Two electrically interactive output ports are included in the differential system of Figure 2.37a. Thus, two Thévenin voltage V_{th1} and V_{th2},

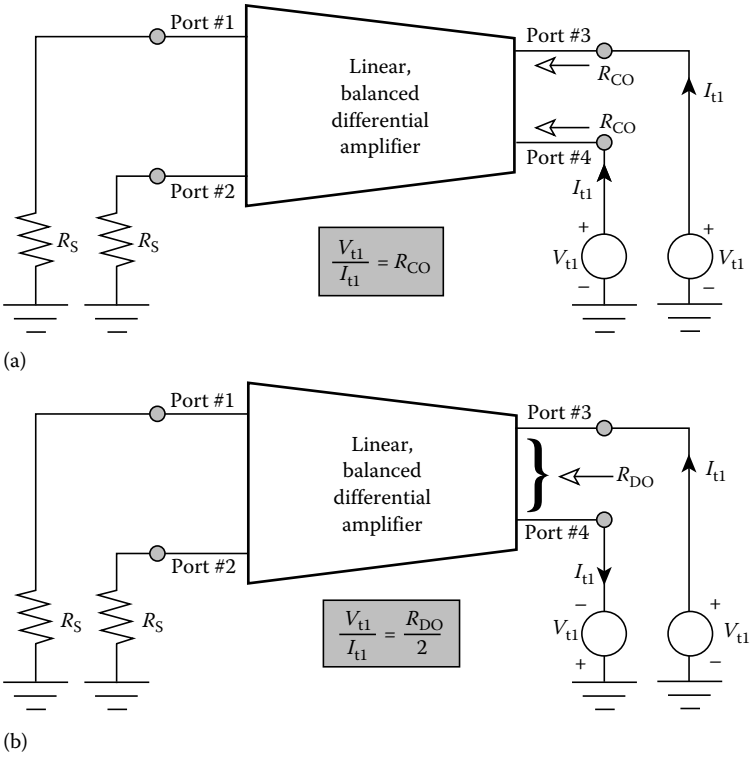


FIGURE 2.43 (a) Test circuit used to evaluate the driving point common-mode output resistance of a linear, balanced differential pair. (b) Test circuit used to evaluate the driving point differential-mode output resistance of a linear, balanced differential pair.

each of which is linearly dependent on the differential-mode and the common-mode components of the applied source signals must be evaluated. These Thévenin output responses derive from open-circuited load conditions, as indicated in Figure 2.44a. With $R_{S1} = R_{S2} = \Delta R_S$, balance prevails, and V_{th1} and V_{th2} are characterized by differential and common-mode components, analogous to the characterization of the terminated outputs, V_{O1} and V_{O2}.

The Thévenin voltages in question derive from an analysis of the equivalent circuit in Figure 2.44b, which represents the model of Figure 2.42 modified to account for nonzero source excitation. The proportionality constants, k_c and k_d, are related to the previously determined common-mode and differential-mode voltage gains. It is a simple matter to confirm that

$$V_{thl},\ V_{th2} = k_c V_{CI} \pm \left(\frac{k_d R_{XO}}{2R_{CO} + R_{XO}}\right)\frac{V_{DI}}{2} \tag{2.107}$$

The first term on the right-hand side of this relationship is the open-circuit common-mode output voltage, while the second term is the open-circuit differential-mode output voltage. It follows that k_c represents the open-circuit common-mode voltage gain, A_{CO}; that is

$$k_c = \lim_{\substack{R_L \to \infty \\ R_{LL} \to \infty}} (A_C) \triangleq A_{CO} \tag{2.108}$$

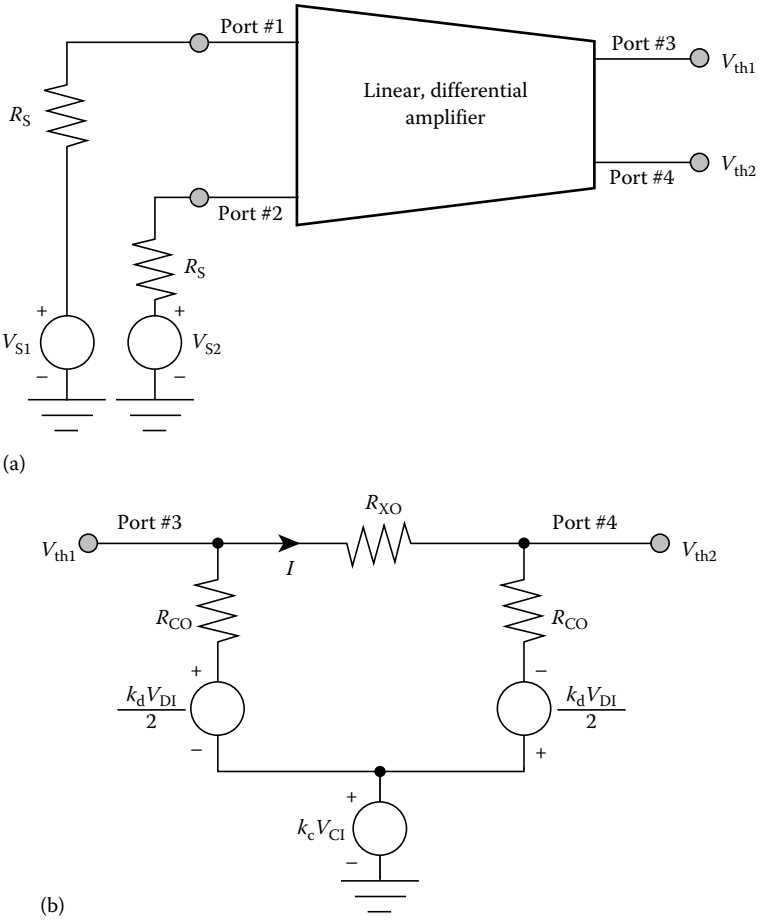


FIGURE 2.44 (a) System schematic diagram used to define the Thévenin voltages at the output ports of a balanced differential amplifier. (b) Thévenin equivalent circuit for the output ports of a balanced differential pair.

On the other hand, the open-circuit differential-mode gain, A_{DO}, is

$$\frac{k_d\ R_{XO}}{2R_{CO} + R_{XO}} = \lim_{\substack{R_L \to \infty \\ R_{LL} \to \infty}} (A_D) \triangleq A_{DO} \tag{2.109}$$

Using Equation 2.53, the Thévenin model parameter, k_d, in the last expression can be cast as

$$k_d = \left(\frac{2R_{CO}}{R_{DO}}\right) A_{DO} \tag{2.110}$$

Figure 2.45 summarizes the foregoing modeling results [13].

Example 2.4

Consider the balanced circuit of Figure 2.46 which is operated as a single-ended input–single-ended output amplifier. The input voltage signal, which is capacitively coupled to the base of transistor Q_1, is

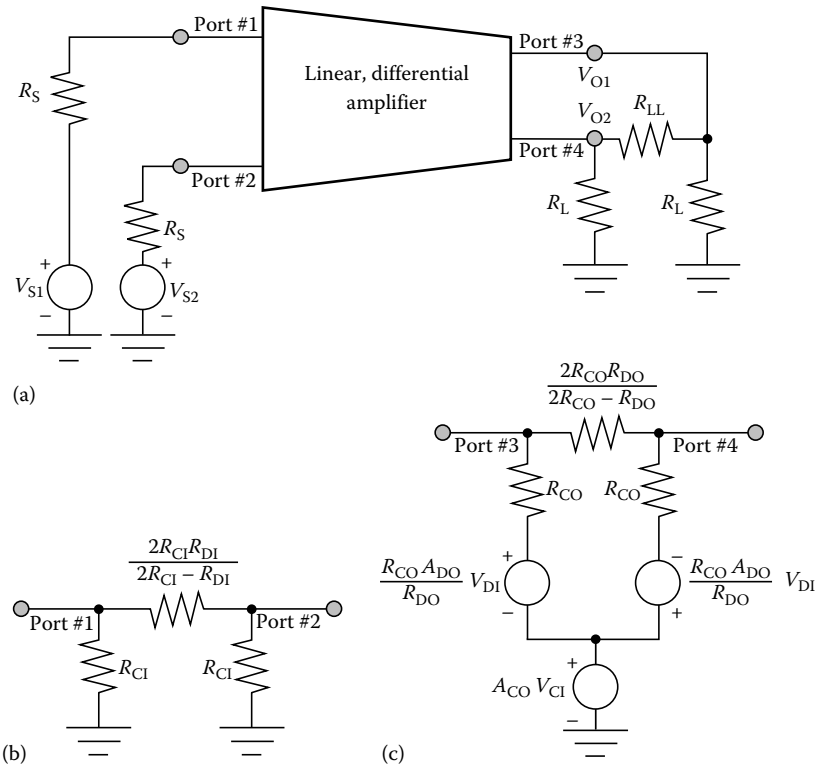


FIGURE 2.45 (a) System schematic diagram of a linear, balanced differential amplifier. (b) Thévenin equivalent input circuit. (c) Thévenin equivalent output circuit. The parameters, A_{DO} and A_{CO}, represent the open-circuit values of the differential- and common-mode voltage gains, respectively, of the balanced pair in (a).

represented as a Thévenin equivalent circuit consisting of the source voltage, V_S, in series with a source resistance, R_S. In order to preserve electrical balance, the base of transistor Q_2 is capacitively returned to ground through a resistance whose value is also equal to R_S. The capacitors can be presumed to act as AC short circuits over the signal frequency range of interest. For the parameters delineated in the inset to Figure 2.46, a computer-aided circuit simulation of the subject amplifier indicates that both transistors have the small-signal parameters $r_b = 33.5\ \Omega$, $r_\pi = 1.22\ k\Omega$, and $\beta = 81.1$. Determine the small-signal voltage gain $A_v = V_{OS}/V_S$, driving point input resistance R_{in}, and driving point output resistance R_{out}.

Solution

1. The AC schematic diagram of the differential-mode half circuit of the balanced amplifier in Figure 2.46 is shown in Figure 2.47a. In concert with earlier arguments, note that the junction of the two emitter degeneration resistances, R_{EE}, is grounded, as are the mid-point of the resistance, R_{LL}, and the node at which R_1, R_2, and the two resistances labeled R are incident. Using the bipolar model of Figure 2.16a, with r_o and r_e ignored, the voltage gain of this structure is the differential-mode voltage gain of the differential pair. Moreover, the driving point input resistance of the circuit at hand is one-half of the differential input resistance of the original pair, while its driving point output resistance is one-half of the differential-mode output resistance. Analysis confirms

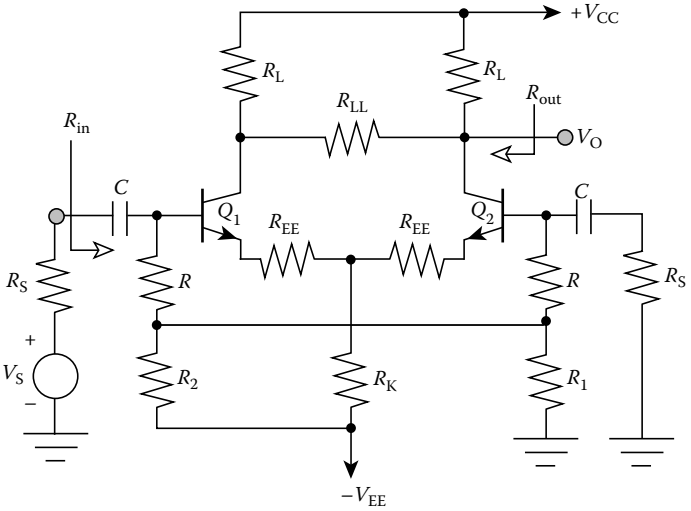


FIGURE 2.46 A balanced bipolar differential amplifier used in a single-ended input–single-ended output mode.

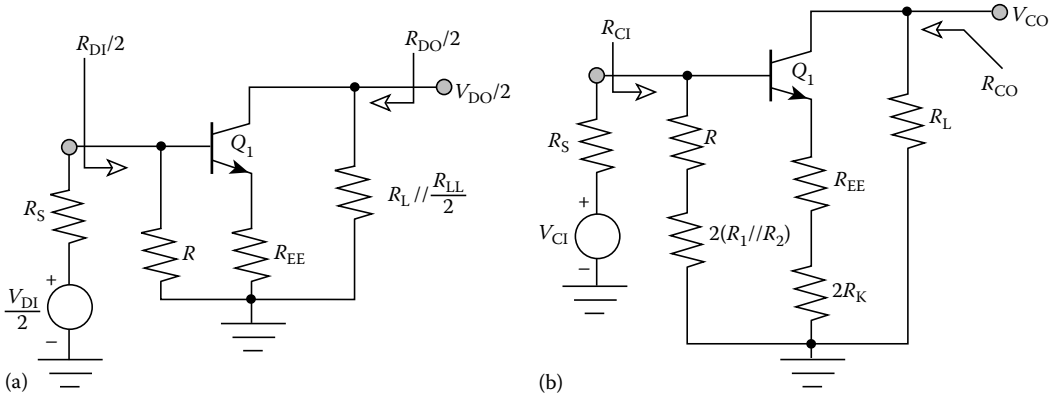


FIGURE 2.47 (a) Differential-mode half-circuit ac equivalent schematic of the differential amplifier shown in Figure 2.46. (b) Common-mode half-circuit ac equivalent schematic of the differential amplifier shown in Figure 2.46.

$$A_D = \frac{V_{DO}/2}{V_{DI}/2} = -\frac{\beta\left(\frac{R}{R+R_S}\right)\left(R_L \| \frac{R_{LL}}{2}\right)}{(R_S \| R) + r_b + r_\pi + (\beta+1)R_{EE}}$$

$$\frac{R_{DI}}{2} = R \| [r_b + r_\pi + (\beta+1)R_{EE}]$$

$$\frac{R_{DO}}{2} = R_L \| \frac{R_{LL}}{2}$$

Numerically, $A_D = -10.09$, $R_{DI} = 5971\ \Omega$, and $R_{DO} = 1875\ \Omega$.

2. The AC schematic diagram of the pertinent common-mode half circuit is given in Figure 2.47b. The input signal voltage is now the common-mode input voltage, V_{CI}, which produces the common-mode output response, V_{CO}. Using the bipolar model of Figure 2.16a, with r_o and r_e ignored, it is easily shown that

$$A_C = \frac{V_{CO}}{V_{CI}} = -\frac{\beta\left(\frac{R+2(R_1\|R_2)}{R+2(R_1\|R_2)+R_S}\right)R_L}{\{R_S\|[R+2(R_1\|R_2)]\}+r_b+r_\pi+(\beta+1)(R_{EE}+2R_K)}$$

$$R_{CI} = [R+2(R_1\|R_2)]\|[r_b+r_\pi+(\beta+1)(R_{EE}+2R_K)]$$

$$R_{CO} = R_L$$

   Numerically, $A_C = -923.3\ (10^{-3})$, $R_{CI} = 5493\ \Omega$, and $R_{CO} = 1500\ \Omega$. The common-mode rejection ratio $p = A_D/A_C$ 10.93 is small owing to the relatively small value of the resistance R_K.
3. As the output voltage is extracted at the collector of transistor Q_2, Equation 2.96b is the applicable equation for determining the output signal voltage, V_{OS}. With only a single source voltage, V_S, applied, the differential input voltage is V_S, and the common-mode input voltage is $V_S/2$. It follows that

$$V_{OS} = A_C V_{CI} - \left(\frac{A_D}{2}\right)V_{DI} = \left(\frac{A_C - A_D}{2}\right)V_S$$

   whence a voltage gain of

$$A_v = \frac{V_{OS}}{V_S} = \frac{A_C - A_D}{2}$$

   These analyses give $A_v = 4.583$.
4. In order to evaluate the driving point input and output resistances, the parameters, R_{XI} and R_{XO} must be calculated. From Equations 2.103 and 2.106, $R_{XI} = 13.08\ \text{k}\Omega$, and $R_{XO} = 5.0\ \Omega$.

The two-port model for calculating the driving point input resistance R_{in} is given in Figure 2.45b. Recall that a circuit resistance, whose value is numerically equal to the internal signal source resistance, R_S, is connected between ground and the node to which the base of transistor Q_2 is incident. By inspection,

$$R_{in} = R_{CI}\|[R_{XI}+(R_{CI}\|R_S)]$$

or $R_{in} = 3.87\ \text{k}\Omega$.

The output port model that emulates the driving point output resistance is given in Figure 2.45c. This model is analogous to that of Figure 2.45a, except that no external loads are connected between signal ground and the node to which the collector of transistor Q_1 is incident. Clearly,

$$R_{out} = R_{CO}\|(R_{XO}+R_{CO})$$

which produces $R_{out} = 1219\ \Omega$.

References

1. J. J. Ebers and J. L. Moll, Large-signal behavior of junction transistors, *Proc. IRE*, 42, 1761–1772, Dec. 1954.
2. H. K. Gummel and H. C. Poon, An integral charge-control model of bipolar transistors, *Bell System Tech. J.*, 49, 115–120, May–June 1970.

3. H. N. Ghosh, A distributed model of the junction transistor and its application in the prediction of the emitter-base diode characteristic, base impedance, and pulse response of the device, *IEEE Trans. Electron Devices*, ED-12, 513–531, Oct. 1965.
4. J. R. Hauser, The effects of distributed base potential on emitter current injection density and effective base resistance for stripe transistor geometries, *IEEE Trans. Electron Devices*, ED-11, 238–242, May 1965.
5. P. R. Gray and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, New York: Wiley, 1977, pp. 16–19.
6. C. T. Kirk, A theory of transistor cut-off frequency (f_T) at high current densities, *IEEE Trans. Electron Devices*, ED-9, 164–174, Mar. 1962.
7. J. M. Early, Effects of space-charge layer widening in junction transistors, *Proc. IRE*, 46, 1141–1152, Nov. 1952.
8. A. S. Sedra and K. C. Smith, *Microelectronic Circuits*, New York: Holt, Rinehart & Winston, 1987, pp. 52–57, 639–642.
9. A. B. Grebene, *Bipolar and MOS Analog Integrated Circuit Design*, New York: Wiley Interscience, 1984, pp. 170–182.
10. G. R. Wilson, A monolithic junction FET-NPN operational amplifier, *IEEE J. Solid-State Circuits*, SC-3, 341–348, Dec. 1968.
11. E. J. Angelo, *Electronic Circuits*, New York: McGraw-Hill, 1970, Chapter 4.
12. A. B. Grebene, *Bipolar and MOS Analog Integrated Circuit Design*, New York: Wiley Interscience, 1984, pp. 217–224.
13. S. A. Witherspoon and J. Choma, Jr., The analysis of balanced linear differential circuits, *IEEE Trans. Educ.*, 38, 40–50, Feb. 1995.

2.3 MOSFET Biasing Circuits

David G. Haigh, Bill Redman-White, and Rahim Akbari-Dilmaghani

2.3.1 Introduction

CMOS technology is finding a very wide range of applications in analog and analog-digital mixed-mode circuit implementations in addition to its traditional role in digital circuits. In mixed-mode circuits compatibility of analog circuits with digital very large scale integration (VLSI) is important, particularly in cost-sensitive areas and situations where low power consumption is required. Such CMOS circuits require a range of biasing circuits and it is this topic that is the main subject of this section, although the subject is mentioned elsewhere in this text, where particular designs are covered.

The requirements for biasing circuits in CMOS circuit design can be divided into the requirements for voltage and for current sources. These sources can be further subdivided into two additional categories, high precision and noncritical. High-precision voltage or current sources are essential components in data converters, both analog-to-digital and digital-to-analog and the precision required depends on the overall target precision of the data converter. For a large number of bits, the precision required could be very great indeed and would need to be maintained over a specified temperature and supply voltage range and in the presence of on-chip, chip-to-chip and wafer-to-wafer component parameter variations. Precision sources are also required in other applications such as dc pedestals for video signals in video systems.

Noncritical voltage sources are generally required for setting up an internal analog ground or for biasing the gates of FETs in common-gate configuration, as in a cascode FET. In these cases, the sensitivity of overall circuit performance parameters to the bias voltage would generally not be high and moderate precision circuit techniques would be acceptable. The main considerations would be to

maintain correct operation, especially in terms of signal headroom, overall variations in process, power supply, and temperature conditions. The requirement for current sources for biasing CMOS analog and mixed-mode circuits is very considerable. The reason for this is that most circuits, such as an operational amplifier, consist of several stages, each of which requires biasing. In discrete circuits, the tendency is to use resistors for biasing, for reasons of cost and the poor sample-to-sample tolerances on discrete active device parameters. In integrated circuit implementation, on the other hand, relatively well-matched devices on the same chip are available and the use of current source biasing minimizes gain loss due to loading effects. Furthermore, using a multioutput current mirror to supply different parts of the circuit allows stabilization of the current against temperature and power supply voltage variations to be performed at one location only (on or off the chip) and the stabilized current can be distributed throughout the chip or subcircuit using current mirror circuits. This also produces significant immunity to localized power supply fluctuation and noise. Full or partial stabilization of bias currents with operating and environmental changes is desirable in order to minimize the range of operating current for which design must be specified, allowing higher design performance targets to be achieved.

In many cases, CMOS circuits have their power supplies derived from an off-chip bipolar regulator (with its own internal bandgap). In these situations, a reasonable voltage reference can sometimes be obtained from the power supply voltage via a potential divider. A voltage obtained in this way can be applied to an external low tolerance resistor to obtain a current reference of moderate precision. The value of realizing a reference on chip is that the cost of the external reference can be avoided. For example, in battery supplied equipment, a large degree of supply voltage immunity is required in the presence of a widely varying battery voltage. It is possible to use an on-chip voltage regulator or an on-chip, switched-mode power supply. In CMOS technology, high-precision voltage references are difficult to design, although a reference with good power supply rejection is possible. BiCMOS technology overcomes many of the problems experienced with CMOS technology since well-controlled bipolar devices for very high-precision references are available on-chip.

Many references to biasing appear in this text under the heading of the circuit concerned and only some general guidelines and principles, together with some example circuits will be given here. We begin this section on CMOS biasing circuits by considering the devices available for biasing in a CMOS-integrated circuit including parasitically realized bipolar junction devices. Some useful simplified models of these devices and a brief examination of the variability of the relevant model parameters will be presented. We then consider different types of references and biasing circuits. Since voltage and current references are closely interrelated, they are dealt with concurrently. The material is presented according to a gradually increasing level of sophistication and achievable precision, starting with simple circuits with only minor supply voltage, temperature, and process independence and leading to fully curvature-compensated bandgap references. This is followed by a consideration of references based on less usual devices that may not be available or usable in every process but that offer potentially attractive solutions. We then illustrate the application of some biasing techniques in the context of simple operational amplifier circuits. Finally, we consider the biasing of amplifiers for very low supply voltages, where rail-to-rail optimized performance is required, and dynamic biasing techniques. The topic of CMOS biasing circuits is sufficiently large to warrant an entire book and we include a list of references that will help provide the reader with more detailed information.

2.3.2 Device Types and Models for Biasing

2.3.2.1 Devices

The principal devices of CMOS technology are the enhancement-mode N-channel and P-channel MOSFET, which are shown schematically in Figure 2.48a and b for N-well and P-well technologies, respectively. Currently, N-well technology is more widely available than P-well. Depletion-mode devices

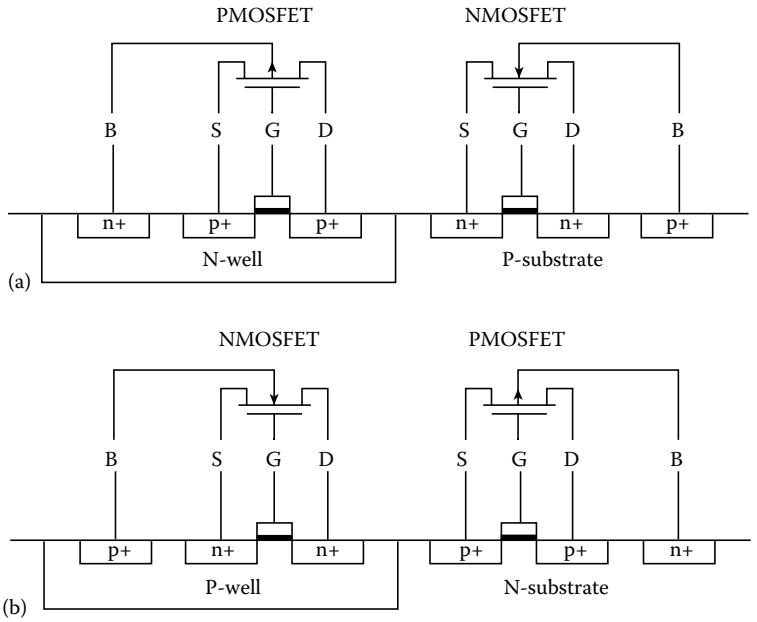


FIGURE 2.48 Realization of NMOS and PMOS FETs in CMOS technology. (a) N-well process. (b) P-well process.

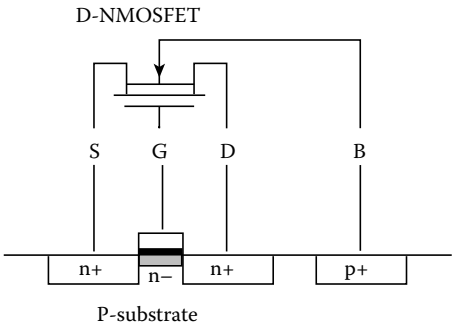


FIGURE 2.49 Realization of NMOS depletion-mode FET in N-well CMOS technology.

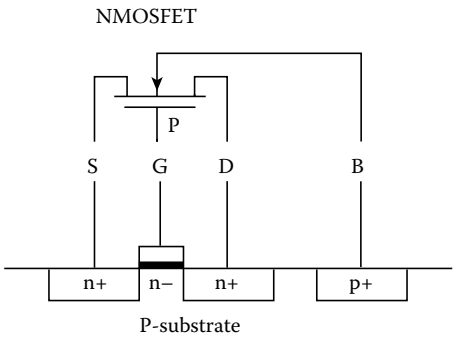


FIGURE 2.50 NMOS enhancement-mode FET with P-doped polysilicon gate.

are not routinely provided in CMOS (in contrast with NMOS), but they are available in some processes [23] and they can be used to realize reference circuits [4,5,23]. They are produced by an additional implementation under the gate region (N-type for NMOS and P-type for PMOS). The NMOS depletion-mode symbol is shown schematically in Figure 2.49.

In most processes, the MOSFET gate material is highly N-doped polysilicon. In some processes [8], P-doped gates are available, and these can be used to realize reference circuits. The symbol of an N-channel enhancement MOSFET with P-type doped base is shown in Figure 2.50.

In order to realize temperature and process desensitized biasing of CMOS circuits, the special properties of BJTs are advantageous. BJT devices can be realized in CMOS technology as parasitic devices with certain restrictions. Two classes of such device are available, namely, vertical and lateral. For the vertical device [13], the restrictions are that an N-well process can realize PNP devices with the collectors connected to the substrate (most negative supply rail) and that a P-well process can realize

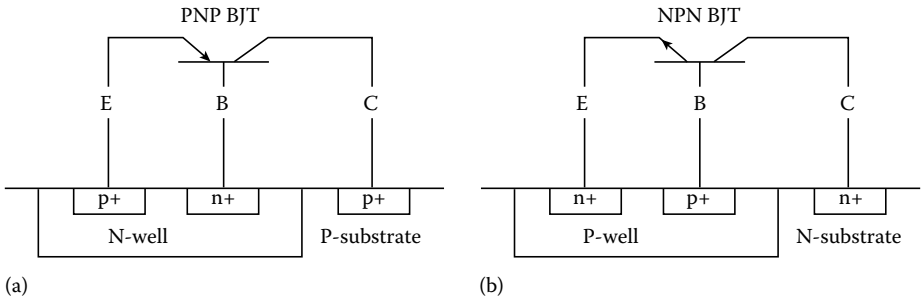


FIGURE 2.51 Realization of vertical BJT devices in CMOS technology. (a) N-well process. (b) P-well process.

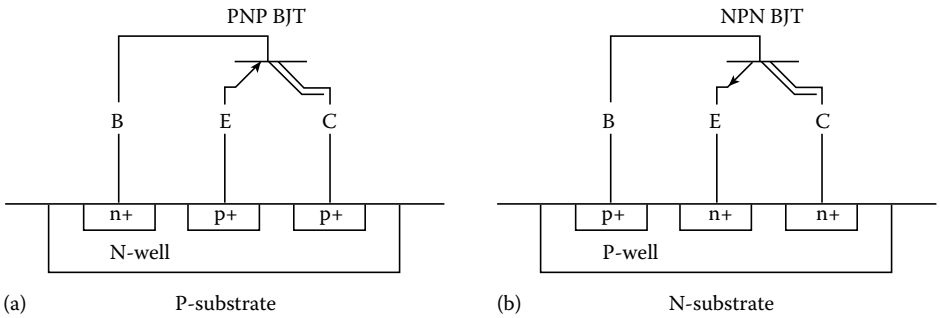


FIGURE 2.52 Realization of lateral BJT devices in CMOS technology. (a) N-well process. (b) P-well process.

NPN devices with the collectors connected to the substrate (most positive supply rail). The realization of these vertical devices in the case of N-well and P-well processes is shown schematically in Figure 2.51a and b, respectively. The devices can have typical current gains of around 100 and may have high leakage, and certain precautions have to be taken in the layout. One problem is that the control on the parameters of these devices in production is minimal. Nevertheless, such devices are adequate to realize moderate-to-high precision bias and reference circuits.

The restriction on the collector connections of the vertical BJT devices in CMOS technology is effectively removed in lateral BJT devices [16]. The realization of these devices in the case of N-well and P-well processes is illustrated in Figure 2.52a and b, respectively. It should be noted that β_∂ for lateral bipolar transistors is not related to α_∂ in the usual way due to substrate currents. For both vertical and lateral parasitic bipolar devices, it is often the case that the foundry will not provide detailed characterization and models, and also that the parameters of the devices will not be well controlled.

Apart from MOSFETs and BJTs, the remaining component needed to realize bias circuits are resistors and in some cases capacitors. In the case of precision bias circuits, the realized variable (voltage or current) would be dependent on resistor ratios rather than on absolute values. On-chip resistors may be realized using polysilicon, or as N- or P-well diffusion, as illustrated in Figure 2.53. Diffused resistors are sensitive to substrate potential and thus they are not suitable for precision potential dividers. In some cases, voltages or currents on a chip that are required to have high stability are referred to an off-chip highly stable and accurate discrete resistor. In some advanced processes, film resistors (usually nichrome) are available and they have excellent temperature stability, a wide range of values and can sometimes be laser-trimmed.

Capacitors, both on-chip and external, are used for decoupling bias and reference voltages and are especially valuable where low noise is critically important. On-chip capacitors also provide the basic components used for dynamic biasing.

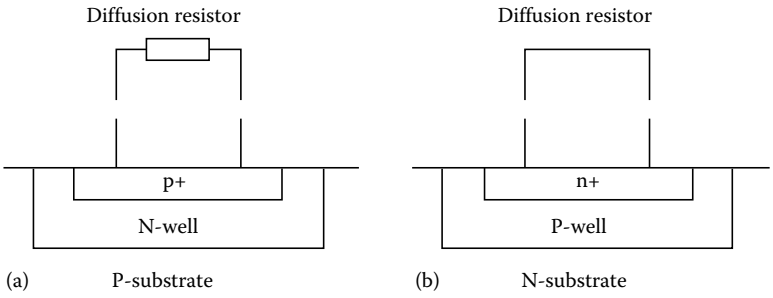


FIGURE 2.53 Realization of diffusion resistors in CMOS technology. (a) N-well process. (b) P-well process.

2.3.2.2 Device Models and Parameter Variability

The MOSFETs in Figure 2.48 may be very approximately described by [13,15]

$$I_d = \beta(V_{gs} - V_t)^2(1 + \lambda V_{ds}) \tag{2.111}$$

where

$$\beta = \mu C_{ox} W/L \tag{2.112}$$

and

$$V_t = V_{to} + \gamma\left[(2\phi_{fb} - V_{bs})^{0.5} - 2\phi_{fb}^{0.5}\right] \tag{2.113}$$

with

$$\phi_{fb} = \left(\frac{kT}{q}\right)\ln\left(\frac{N_{sub}}{n_i}\right) \tag{2.114}$$

and

$$\mu \propto T^{-\eta} \tag{2.115}$$

The parameters in Equations 2.111 through 2.115 are defined in Table 2.1. As a result of Equations 2.111 through 2.115, MOSFET parameters show considerable temperature dependence. In particular, threshold voltage varies with temperature according to

$$\frac{\partial V_t}{\partial T} = \frac{2}{T}\left(\phi_F - \frac{E_{go}}{2q}\right)\left(1 + \frac{\gamma}{2(2\phi_{fb} - V_{bs})}\right) \tag{2.116}$$

which amounts typically to about −2 mV/°C (Ego is the bandgap of silicon at 0 K) [15]. Transconductance varies according to

$$\frac{\partial g_m}{\partial T} = \frac{g_m}{2}\left(-\frac{\eta}{T} + \frac{1}{I_d}\frac{\partial I_d}{\partial T}\right) \tag{2.117}$$

The temperature and process dependencies of MOS devices have led to the exploitation of some properties of combinations of less usual MOS devices. It has been observed that the difference between

TABLE 2.1 Device Parameters

Symbols	Parameters
I_d	Drain current
β	Transconductance parameter
V_{gs}	Gate–source voltage
V_t	Threshold voltage
λ	Output conductance parameter
V_{ds}	Drain–source voltage
μ	Mobility
C_{ox}	Oxide capacitance per unit area
W	Gate width
L	Gate length
V_{to}	Zero substrate bias threshold voltage
γ	Body factor
ϕ_{fb}	Fermi potential
V_{bs}	Substrate–source voltage
K	Boltzmann's constant
T	Absolute temperature in K
Q	Electronic charge
N_{sub}	Substrate doping density
n_i	Intrinsic carrier density
$\acute{\eta}$	Mobility temperature coefficient
V_{FB}	Flat-band voltage
Q_{SS}	Surface charge per unit area
ϕ_P	Bulk potential
Q_d	Charge per unit area in inversion layer
ϕ_{bi}	Channel to substrate built-in potential
Q_i	Implanted charge per unit area
C_{impl}	Capacitance defined by implanted channel depth
ϕ_G	Bandgap voltage for silicon
ϕ_{GO}	Bandgap voltage for silicon at 0 K

the threshold voltages of an enhancement- and depletion-mode FET pair is relatively temperature independent. The threshold voltages for enhancement- and depletion-mode MOSFETs may be written

$$V_{t(enh)} = V_{FB} - \frac{Q_{SS}}{C_{ox}} + 2|\phi_P| + \frac{|Q_d|}{C_{ox}} \tag{2.118}$$

$$V_{t(depl)} = V_{FB} - \frac{Q_{SS}}{C_{ox}} + \phi_{bi} + (|Q_d| - |Q_i|)\left(\frac{1}{C_{ox}} + \frac{1}{C_{impl}}\right) \tag{2.119}$$

where the meaning of the parameters is given in Table 2.1 [4]. Many of the parameters in Equations 2.118 and 2.119 show considerable temperature dependence. The difference between the threshold voltages is given by

$$\begin{aligned} V_{t(diff)} &= V_{t(enh)} - V_{t(depl)} \\ &= 2|\phi_P| - \phi_{bi} - |Q_i|\left(\frac{1}{C_{ox}} + \frac{1}{C_{impl}}\right) \end{aligned} \tag{2.120}$$

assuming that $1/C_{ox} \gg 1/C_{impl}$, which is true in practice. In practice, it is also the case that $2|\phi_{P|} \approx \phi_{bi}$. Since the implanted charge Q_i is controllable and independent of temperature to first order [4], the threshold voltage difference exhibits temperature independence to first order.

The difference between the threshold voltages of two N-channel FETs with polysilicon gates of opposite doping polarity (P and N) also shows relative insensitivity to temperature variations [8]. To a first approximation, the threshold voltage difference is given by

$$\Delta V_G = \phi_G = 1.12 \text{ V (room temperature)} \tag{2.121}$$

which is the bandgap voltage for silicon [8]. A more detailed analysis [1] gives

$$\Delta V_G(T) = \phi_{GO} - \frac{\alpha T^2}{T + \beta} \tag{2.122}$$

where $\alpha = 7.02 \times 10^{-4}$ V/K, $\beta = 1109$ K and the meaning of the remaining parameters is given in Table 2.1. In practice, the degree of temperature independence obtained provides useful reference circuits [8]. The exploitation of both the enhancement–depletion FET threshold difference and the N–P-doped polysilicon gate threshold voltage difference for the design of references will be described later.

The strong temperature dependence of conventional MOS device parameters means that for stable biasing circuits, MOS devices are mainly useful where the critical variable depends on a ratio of parameters of similar devices. Even in this case, the matching is not as good as for bipolar devices. The matching of the gate–source voltages of two similar devices with nominally identical drain currents is inversely proportional to the square root of the gate area and is typically of the order of 10 mV, which limits the minimum offset voltage of a CMOS op-amp. Since op-amps form key components in many voltage and current reference circuits, this is a serious limitation.

In contrast to the rather complex dependence of MOS device parameters with temperature, the situation in the case of BJT devices is relatively straightforward [15,17,26]. The BJT may be described by

$$I_c = I_s e^{V_{be} q/kT} \tag{2.123}$$

where the additional parameters are defined in Table 2.1. Equation 2.123 may alternatively be written

$$V_{be} = \frac{kT}{q} \ln \frac{I_c}{I_s} \tag{2.124}$$

For two devices with an emitter area ratio of A

$$A = \frac{I_{s1}}{I_{s2}} \tag{2.125}$$

We have

$$\begin{aligned} \Delta V_{be} &= V_{be1} - V_{be2} \\ &= \frac{kT}{q} \ln \frac{I_{c1}}{I_{c2}} \frac{1}{A} \\ &= V_T \ln \frac{1}{A} \quad (\text{for } I_{c1} = I_{c2}) \end{aligned} \tag{2.126}$$

Thus, the difference between the V_{be}s of two BJTs with different current densities is proportional to the thermal voltage V_T, which is proportional to absolute temperature (PTAT). The positive temperature coefficient of V_T can be effectively used to cancel the negative temperature coefficient of V_{be} [13]. This is referred to as the bandgap principle.

Resistors are also key elements in MOS biasing circuits and they may be realized using diffusion, polysilicon and, in some advanced processes, using film techniques. Polysilicon and diffused resistors suffer from a high temperature coefficient that is positive for diffusion. The resistivity of gate polysilicon is typically rather low at about 20 Ω/square and its initial value tolerance is quite high. Film resistors have a very low temperature coefficient.

2.3.3 Voltage and Current Reference and Bias Circuits

2.3.3.1 Supply-Voltage-Referenced Voltage and Current References

When the supply voltages to a chip are well-regulated off-chip, then a voltage reference, acceptable in some cases, can be realized by a simple potential divider from the power supply voltage, as shown in Figure 2.54. An external decoupling capacitor may be used if needed and the voltage dividing elements may be resistors (Figure 2.54a) or MOSFETs (Figure 2.54b). If the power supply voltages are well-controlled off-chip, using an external regulator circuit, then a simple current reference can be realized by the arrangement in Figure 2.55, where the reference current is defined by applying a well-defined fraction of the controlled supply voltage to a well-controlled external resistor.

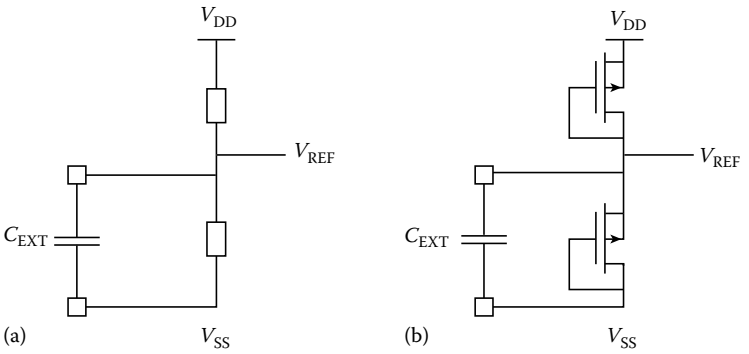


FIGURE 2.54 Voltage reference obtained via potential divider from regulated power supply: (a) using resistors; (b) using MOSFETs.

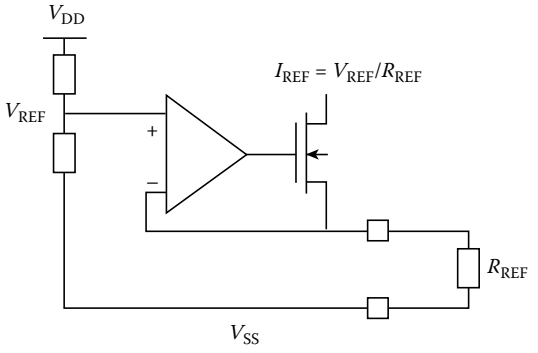


FIGURE 2.55 Current source realized using potential divider.

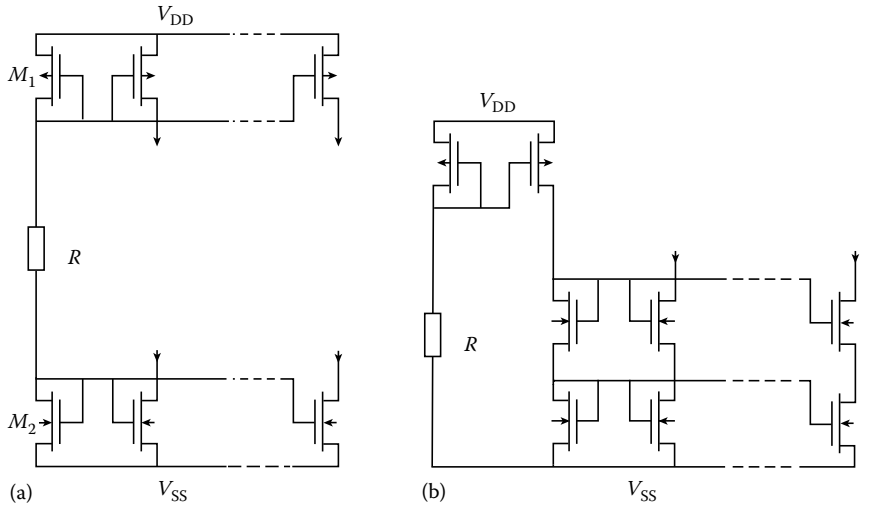


FIGURE 2.56 Resistor/current mirror bias circuits: (a) simple; (b) cascode.

A very simple biasing circuit proving multiple current sources and sinks, as required in CMOS analog signal-processing circuits, is shown in Figure 2.56a [13]. Assuming large FET W/L ratios, the voltage across the resistor is approximately $V_{dd} - V_{ss} - 2V_t$. The current in R is mirrored in the output MOSFETs where the W/L ratios may be chosen to provide required current magnitudes. The resistor R may have to be realized off-chip as a precision film component with narrow tolerance and small temperature coefficient. In practice, the voltage across the diode-connected MOSFETs M_1 and M_2 will be greater than V_t and have some dependence on MOSFET β as well as V_t.

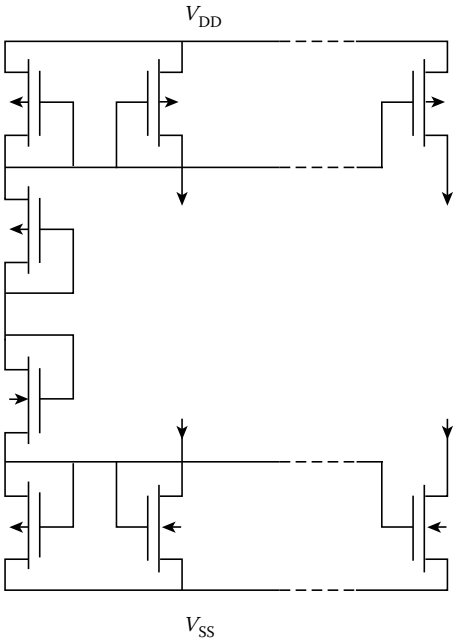


FIGURE 2.57 FET/current mirror bias circuit.

For the circuit in Figure 2.56a, the source conductance is equal to the MOSFET g_{ds}, which might not be sufficiently low for some applications. This disadvantage can be overcome by introducing cascode FETs as in Figure 2.56b. Nevertheless, the supply voltage dependence of the circuit in Figure 2.56 remains.

For uncritical applications where the current source is to be realized entirely on-chip, the resistor R in Figure 2.56a may be replaced by a chain of diode-connected P- and N-channel MOSFETs, as shown in Figure 2.57 [13]. The number of these devices and their W/L ratios may be chosen according to the supply voltage and the value of the current to be realized. Since the diode-connected MOSFETs are effectively realizing the resistor in the basic current source of Figure 2.56a, the power supply voltage dependence of the current remains. In addition, the effective resistance realized depends on MOSFET β and V_t, both of which have large tolerances and high temperature coefficients.

2.3.3.2 MOSFET Threshold Voltage-Based References

A current reference with reduced power supply voltage dependence is shown in Figure 2.58 [13]. By choosing W/L for M_1 to be large, the gate–source voltage of M_1 can be made close to the device threshold voltage V_t. Since the gate–source voltage of M_1 appears across the resistor R, the current in R is approximately V_t/R, which is ideally independent of power supply voltage. The W/L ratios of MOSFETs M_3 and M_4 are chosen to define a fixed ratio for the currents in M_1 and R. The combination of MOSFETs M_2, M_3, and M_4 constitute a positive feedback loop and it is important to choose the W/L ratios so that the loop gain is less than unity to avoid oscillation. Many reference circuits have a stable state with all currents zero. In such cases, it is necessary to provide a start-up circuit [13] to prevent the reference circuit locking into an undesired operating point. A source follower can be introduced at the gate of M_2 such that in this condition a current is injected into the circuit. The added components must be such that in the normal operating point, they are switched off and therefore do not influence operation. In practice, the currents realized by the circuit in Figure 2.58 will have some supply voltage dependence due to channel length modulation in the MOSFETs. This effect can be reduced by introducing cascode devices appropriately. Although the currents realized can be made substantially independent of supply voltage, the dependence on resistance R remains. For high precision and temperature independence, R may need to be realized as an off-chip film resistor. It must be borne in mind that the device threshold voltage on which the current depends is rather variable (typically 0.5–0.8 V) and also rather temperature dependent. Solution of this problem requires the introduction of alternative techniques based on BJT or unconventional CMOS devices, which will be discussed.

An alternative circuit to that in Figure 2.58 is shown in Figure 2.59 [25]. This circuit regulates the MOSFET drain currents with the result that MOSFET transconductance is proportional to $1/R$. This circuit also relies on positive feedback and care must be taken in the design to avoid instability. It has been shown in Ref. [21] that a practical stable design results from the choice $(W/L)_4 = (W/L)_3$ and $(W/L)_2 = 4(W/L)_1$, giving $g_m = 1/R$. In processes where R can be realized as a film resistor on-chip, this circuit can stabilize transconductances to within 3% over a 100°C temperature range [25].

2.3.3.3 BJT V_{be}-Based References

The problem that MOSFET threshold voltage is not very well controlled from chip sample to chip sample leads to the idea of using the V_{be} of a parasitic bipolar transistor [13]. Such a V_{be}-referenced circuit is

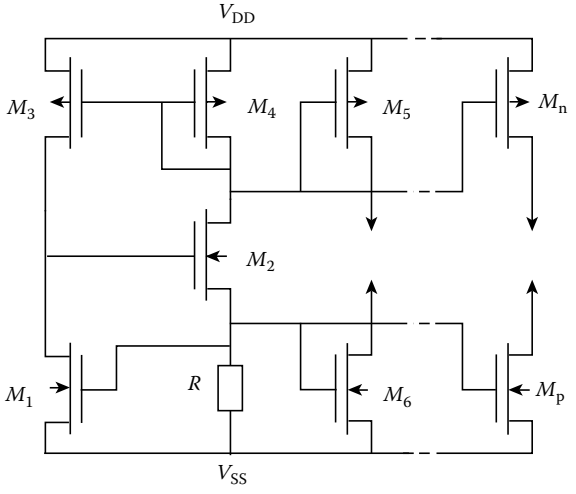


FIGURE 2.58 V_t-referenced current bias circuit.

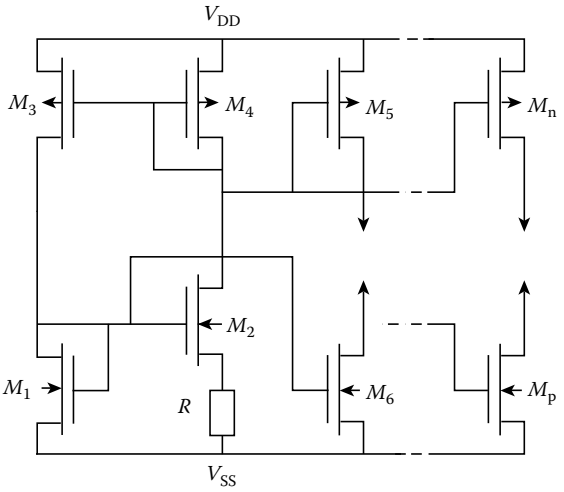


FIGURE 2.59 g_m-R tracking current reference circuit.

shown in Figure 2.60 for the case of an N-well process, where the BJT is PNP [13]. The W/L ratios for M_1 and M_2 are made large so that the V_{be} of T_1 appears substantially across R. In order to achieve high precision and temperature independence, R would need to be an off-chip film resistor. However, the V_{be} has a process-dependent tolerance of about 5% and a dependence with temperature of about -2 mV/°C.

2.3.3.4 BJT V_T-Based References

The temperature dependence of V_{be} in the circuit of Figure 2.60 can be overcome in the V_T-based circuit of Figure 2.61 [13]. The emitter areas of the BJTs Q_1 and Q_2 are scaled in the ratio 1:n and the MOS current mirrors force the emitter currents to be equal. The difference between the V_{be} of Q_1 and Q_2 given by Equation 2.121 appears across the resistor R, hence defining the current. The positive temperature coefficient of V_T can be used to counteract the positive temperature coefficient of the resistor R to obtain

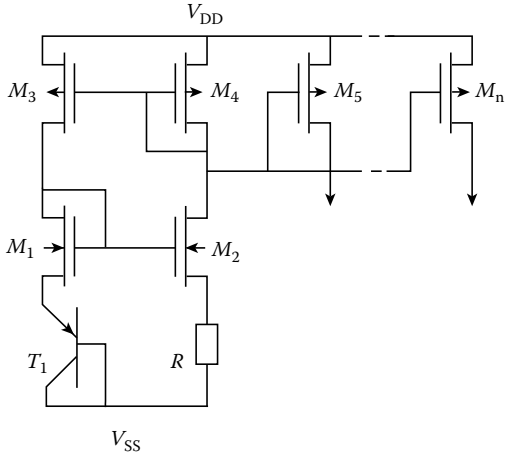


FIGURE 2.60 V_{be}-referenced current bias circuit.

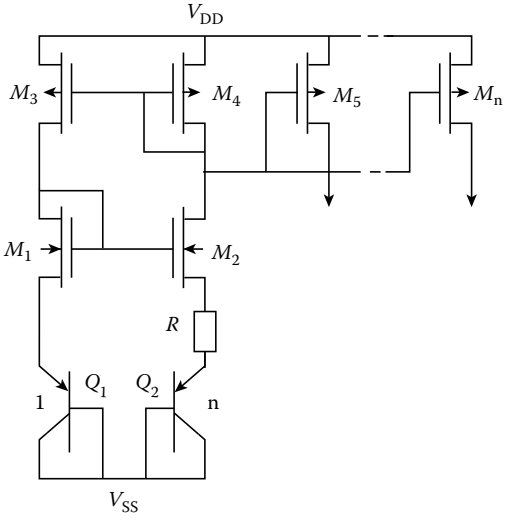


FIGURE 2.61 V_T-referenced current bias circuit.

a stable current. In the circuit of Figure 2.61, M_1 and M_2 must have large W/L to minimize the effect of MOSFET process variability. Also, cascoding of the current mirrors may be required to reduce the effect of channel length modulation.

2.3.3.5 Bandgap References

Precision voltage sources are key requirements for the realization of precision data converters and have received much attention [2,3,7,10,12–14,20,22]. The requirements for high precision and very low temperature and supply voltage dependence have led to the development of the bandgap principle [2,3,7,10,13]. The bandgap principle was originally developed for bipolar technology and a typical architecture is shown in Figure 2.62a. As described previously, the difference between the V_{be}s of two BJTs with different current densities is proportional to the thermal voltage V_T and is PTAT. In Figure 2.62a, the difference between the V_{be}s appears across R_3 and in scaled form across R_2. Thus the output voltage is equal to the V_{be} of Q_1 plus the scaled version of ΔV_{be}. Thus R_2/R_3 may be chosen so that the opposite temperature coefficients of V_{be} and ΔV_{be} cancel. The ratio R_1/R_2 determines the ratio of the currents in Q_1 and Q_2. The circuit in Figure 2.62a is incompatible with implementation using CMOS technology with vertical parasitic bipolar devices because the collectors are not grounded. This can be overcome using the architecture in Figure 2.62b. However, there is the further problem that the offset voltage of the operational amplifier is multiplied by the internal gain of the feedback loop and added to the output. Offset is worse for CMOS than for bipolar operational amplifiers. This problem has been overcome in various ways. In Refs. [12,22], use is made of a discrete time offset compensated differential amplifier, which can have very low offset. Another approach is to make use of lateral bipolar devices, which do not suffer from the topological restrictions of their vertical counterparts [16]. Thus the architecture of Figure 2.62a or an equivalent topology may be implemented.

A typical example of a current reference based on a bandgap voltage reference is shown in Figure 2.63 [13]. The current in the resistor xR is V_T-referenced as in Figure 2.61 and therefore has a negative temperature coefficient (the BJTs are vertical parasitic devices). This current is converted to a voltage and weighted by the resistor xR before being added to the V_{be} of Q_3, which has a negative temperature coefficient. The parameter x is chosen to obtain an overall zero temperature coefficient for the output current, which is given by V_{REF}/R_0. Clearly, R_0 needs to be a high-precision resistor and could be external to the chip. The current mirrors need to be very well matched as any offset is amplified. The operational amplifier needs to have a low offset voltage since this is added to the reference voltage. Further current mirroring may be used to change the sign of the current or to increase the permissible range of the output voltage, referred to as compliance.

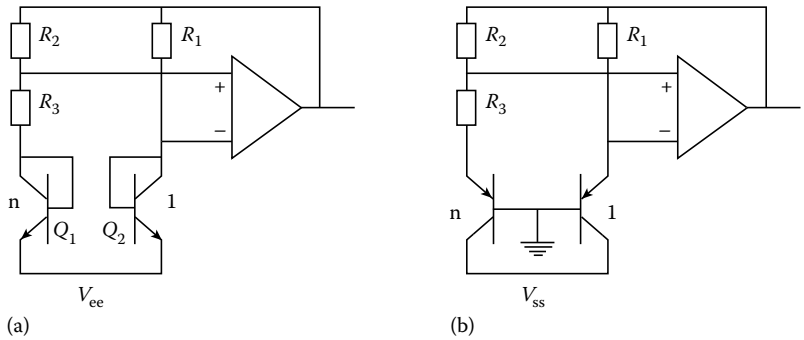


FIGURE 2.62 Basic bandgap circuits. (a) Classical bandgap circuit. (b) Modified form with grounded-collector PNP transistors, assuming an N-well process.

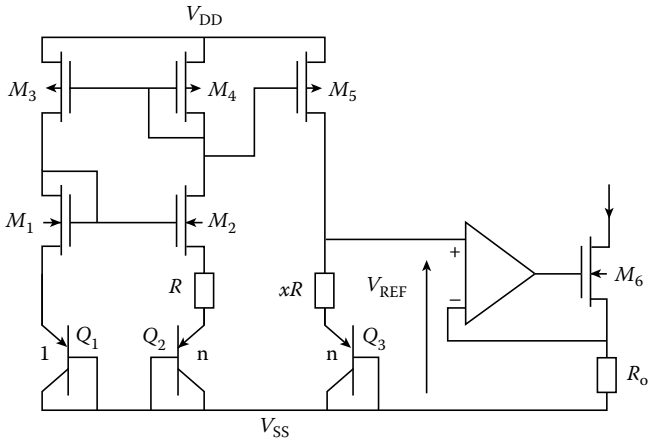


FIGURE 2.63 Bandgap current bias circuit.

2.3.3.6 Curvature-Compensated Bandgap References

The bandgap reference principle can provide a zero temperature coefficient at a single temperature, leaving a temperature dependence that is dominated by a second-order temperature dependence. Very sophisticated techniques have been developed [12,24] to eliminate this second-order dependence to leave a typically much smaller third-order dependence. This technique is referred to as curvature compensation. An example of a curvature-compensated current reference [24] is shown in Figure 2.64. This circuit can achieve precisions of the order of 5 ppm/°C for supply voltages over 5–15 V.

2.3.3.7 Discrete Time Bandgap References

The voltage reference in Ref. [12] provides curvature compensation and achieves a drift of the order of 13 ppm/°C over the commercial temperature range. The design is based on a comprehensive analysis of nonideal effects in the basic bandgap circuit including finite β and base resistance of the bipolar devices, operational amplifier offset, and bias current variation. This leads to a system involving a very low offset switched capacitor differential amplifier and a system of injecting a differential pair of currents into the

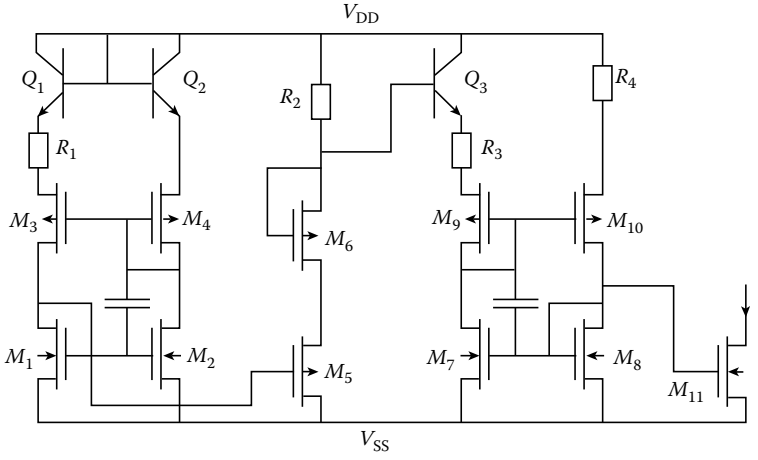


FIGURE 2.64 Curvature-compensated current bias circuit.

emitters of the bipolar devices to provide curvature compensation. The offset cancellation of the switched capacitor differential amplifier is accompanied by techniques for cancellation of the effect of base currents and base resistance in the bipolar devices. Base currents can sometimes be a severe problem due to the available current gains of parasitic bipolar devices. The design is fully compatible with a digital IC process and achieves an equivalent precision of 12 b. Room temperature trims are necessary for a zero temperature coefficient and for curvature compensation. Although low-frequency power supply rejection is good, it falls with increasing frequency.

In Ref. [22], a floating voltage reference for signal-processing applications with a good power supply rejection ratio of at least 85 dB maintained up to 500 MHz is realized. Over a temperature range of −40 to +85°C, voltage dependence is 40 ppm/°C and supply voltage dependence ±5%. The circuit has the important advantage that trimming is not required.

2.3.4 Voltage and Current References Based on Less Usual Devices

2.3.4.1 Use of Device in Subthreshold Region

An alternative approach to current reference making use of MOSFETs in the subthreshold region is reported in Ref. [19]. The principle of the approach is illustrated in Figure 2.65a, where for thermal stabilization the voltage source is required to be PTAT. The PTAT voltage source is realized as a cascade of 5 of the PTAT voltage sources shown in Figure 2.65b, which rely on the subthreshold mode operation of the devices. In practice, cascoding of the current mirrors and current sources is required and a start-up circuit is needed. A current accuracy of 3% with temperature stability of 3% over 0–80°C can be achieved with this approach [19].

2.3.4.2 Voltage Reference Circuits Using Lateral Bipolar Devices

The circuit diagram of a bandgap voltage reference making use of lateral bipolar devices is shown in Figure 2.66 [16]. The circuit is designed to be insensitive to low β and α of the bipolar devices. It is also insensitive to offsets and mismatch. A single trim at room temperature is required and a high power supply rejection ratio, at least at low frequencies, is obtained. The output voltage is stable to within 2 mV over a wide temperature range.

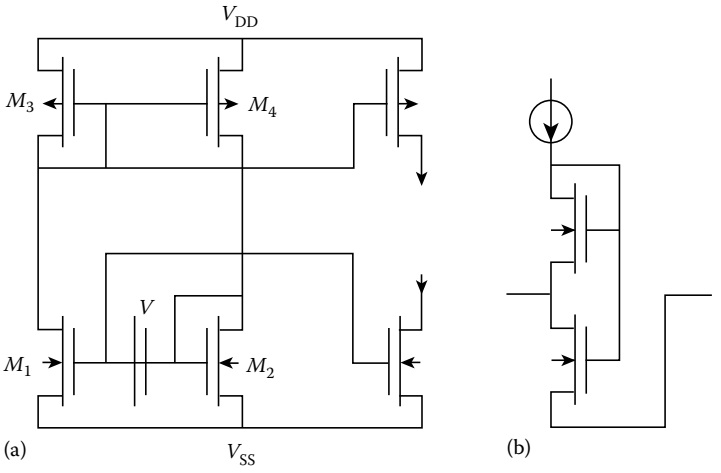


FIGURE 2.65 MOS current bias circuit based on weak inversion operation: (a) basic circuit; (b) voltage source cell.

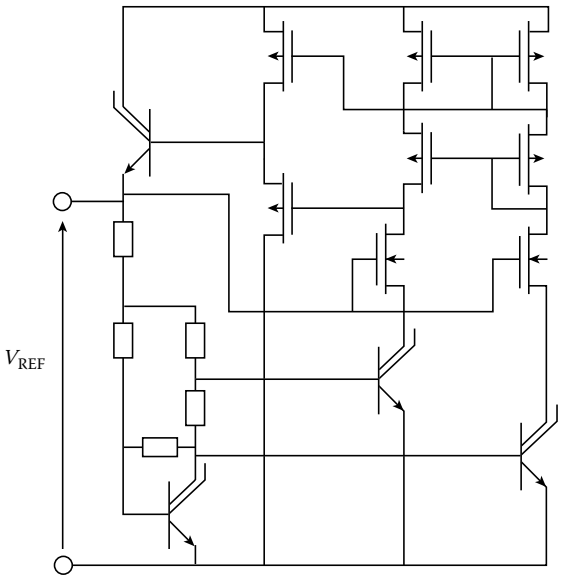


FIGURE 2.66 Voltage reference using lateral bipolar devices.

2.3.4.3 Voltage References Based on Enhancement and Depletion-Mode Threshold Voltage Difference

The topology restrictions and imperfections of the BJT devices available in CMOS technology have led to the development of alternative techniques for designing references without needing bipolar devices. In one technique, the fact that the difference between the threshold voltage of depletion-mode and enhancement-mode devices is relatively temperature independent has been exploited [4,5,23].

The section on device models and parameter variability demonstrated that the difference in threshold voltages of an enhancement- and depletion-mode MOSFET is relatively insensitive to temperature. Since the threshold voltage of the depletion-mode device is negative, this approach leads to a reference voltage of the order of 2 V, which is higher than the bandgap voltage, and this can be an advantage. A basic scheme for exploiting this principle for a voltage reference is shown in Figure 2.67. The op amp adjusts the gate voltage of the enhancement-mode FET to keep the drain voltages the same and the resistor values can be used to adjust the ratio of the currents in the two FETs. The operational amplifier may be implemented at device level [4]. The gate voltage of the depletion-mode FET may be connected to the output of a buffer amplifier whose output voltage may be adjusted using polysilicon fuses to typically 3.15 ± 0.02 V [5].

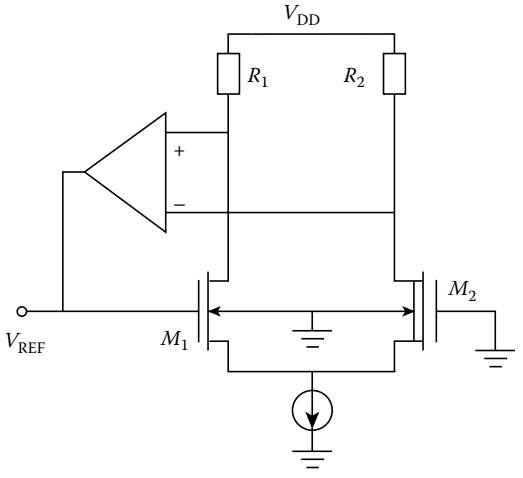


FIGURE 2.67 Basic reference based on enhancement–depletion threshold difference.

Higher reference voltages may be obtained by replicating the enhancement- and depletion-mode MOSFETs. In Figure 2.68, the reference voltage is the difference between the threshold voltages of the three enhancement-mode MOSFETs M_1–M_3 and the three depletion-mode MOSFETs

M_4–M_6 [23]. M_7–M_{10} are providing the necessary bias currents. In Ref. [23], the variation of V_{REF} with temperature is 1.5 mV/°C, which is useful for many biasing situations and the reference voltage is of the order of 3 V in spite of low threshold voltage devices.

2.3.5 Voltage References Based on N- and P-Doped Polysilicon Gate Threshold

2.3.5.1 Voltage Difference

In CMOS technology, the gate material is usually polysilicon with N-type doping. In some processes, selective doping to provide P-type doping of the polysilicon gate is also available and the presence of both types of doping has been exploited for reference circuit design [8].

The basic principle of a voltage reference based on the difference between the threshold voltages of N- and P-doped polysilicon gate MOSFETs is illustrated in Figure 2.69. M_1 has a P-doped gate and a higher threshold voltage than M_2. M_1 and M_2 are in different P-wells but have the same effective dimensions and bias currents.

A full transistor-level implementation of the basic circuit in Figure 2.69 is shown in Figure 2.70 [8]. M_1 and M_2 are the reference MOSFETs. M_3 has a very long channel and its current is the same as that in M_1 by virtue of the current mirror M_4: M_5. Thus the current in M_1 adjusts itself to the crosspoint of the characteristics of M_1 and M_3. M_7, M_8, and M_9 ensure that the currents in M_1 and M_2 are identical. M_6 is a start-up device. When the power supply is switched on, M_6 comes on but within 1 ms is switched off by the reverse leakage resistance of the polysilicon diode D. In Ref. [8], M_1 and M_2 can have a W/L of 100/20 μm; supply voltage sensitivity of $<10^{-3}$ is obtained for V_{DD} between 2 and 9 V [8]. Digital tuning using polysilicon fuses to reference voltages other than the polysilicon gate work function difference can be obtained and a further level of temperature compensation applied [8].

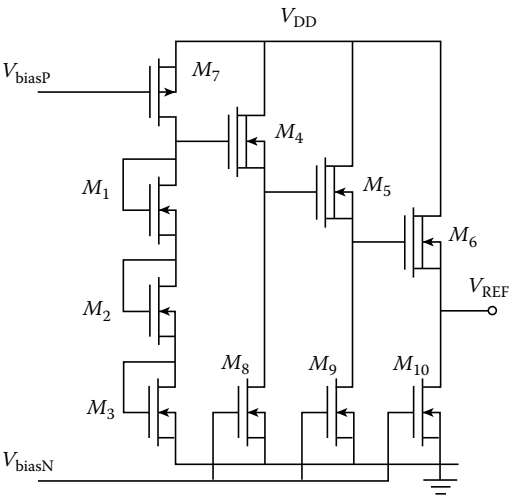


FIGURE 2.68 High-output enhancement–depletion threshold difference reference.

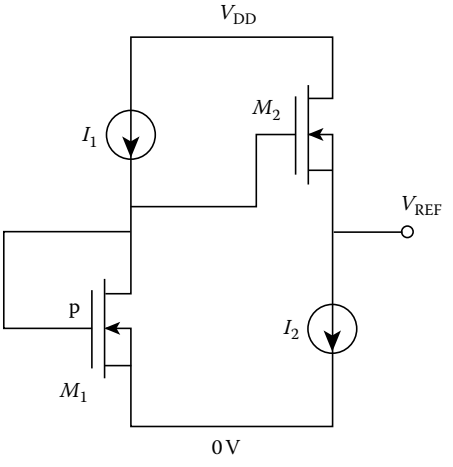


FIGURE 2.69 Basic reference based on polysilicon work function difference.

2.3.6 Biasing of Simple Amplifiers and Other Circuits

2.3.6.1 Simple Amplifiers

In traditional two-stage amplifier design, the bias for the whole circuit is easily set up from one reference current and no critical voltage differences have to be set up. It is only necessary to ensure that the operating currents, and hence the transconductance, of each device have a required value. A typical example of the biasing of a two-stage amplifier is shown in Figure 2.71 [13]. The ratio of currents between

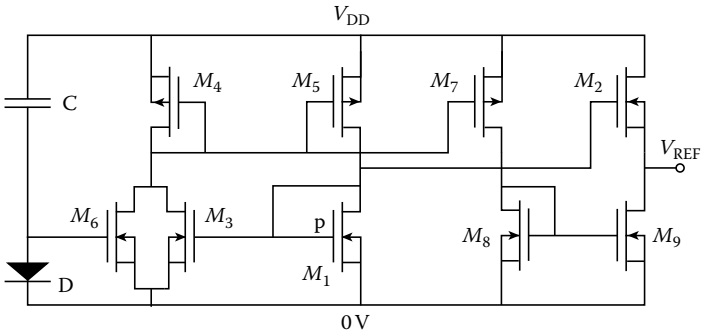


FIGURE 2.70 Example of reference based on polysilicon work function difference.

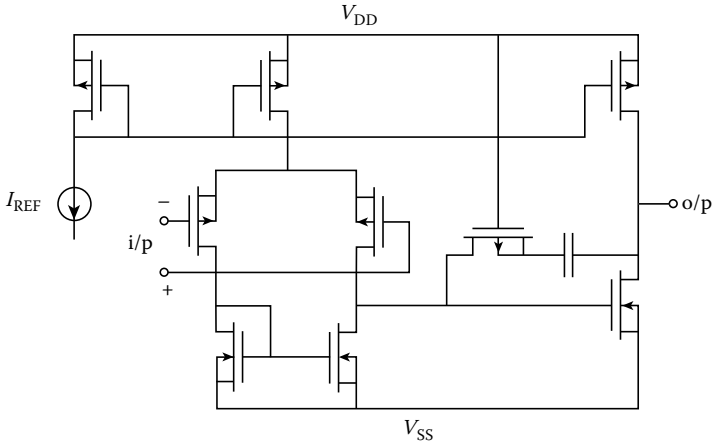


FIGURE 2.71 Two-stage amplifier.

the first- and second-stage controls the separation of the poles and also the systematic offset. The internal biasing circuit consists simply of a set of current mirrors.

2.3.6.2 Cascode Amplifiers

In cascode amplifiers [13], the idea is to raise the amplifier output impedance in order to increase the gain. An important requirement, especially in a low supply voltage environment, is to obtain maximum output voltage swing, or compliance, in the cascode amplifier. This requirement makes the biasing of the cascode devices critical.

A simple amplifier designed for cascode loads is shown in Figure 2.72. Only a single current mirror from the main current reference is needed to control all the bias currents. For reasonable low-frequency gains of say >60 dB, the output impedance must be made high while keeping component parameters practical. Use of very long channel output FETs is undesirable because of poor bandwidth and the chip area requirement. Therefore, the cascode technique, as shown in Figure 2.73, is the ideal solution. This raises output impedance by approximately $g_{m(M11)}r_{ds(M11)}$ and $g_{m(M12)}r_{ds(M12)}$ on each side. However, the maximum output voltage swing, or compliance, of the circuit has now been reduced by at least the saturation voltages of M_{11} and M_{12}. The cascode devices must be biased so that the voltage across M_8 and M_9 is just above V_{dsat}.

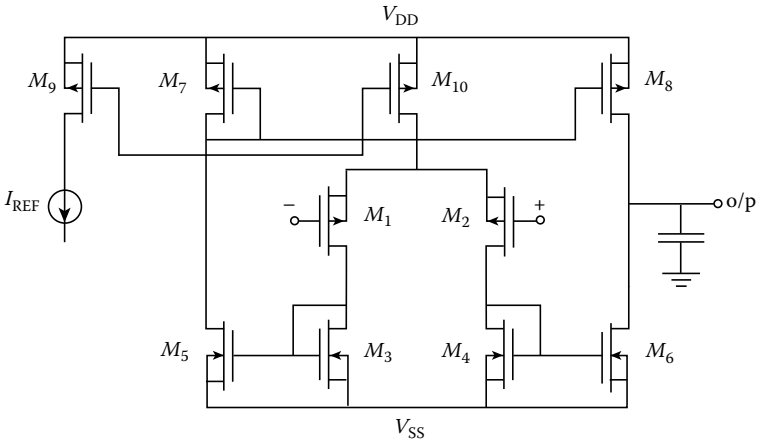


FIGURE 2.72 Simple single-stage amplifier.

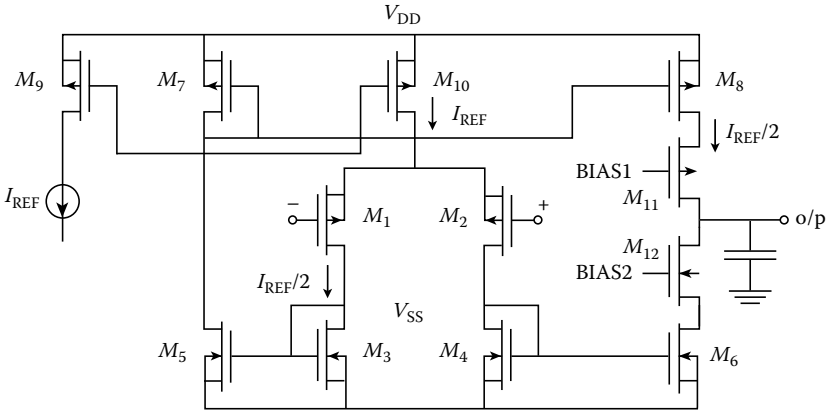


FIGURE 2.73 Simple single-stage amplifier with cascoded output FETs.

The usual way of achieving this is to arrange that a current is passed through an FET with a scaled width so that one obtains a voltage $V_{TN} + V_{Dsat(M6)} + V_{Dsat(M12)}$ for the N-channel side and $V_{TP} + V_{Dsat(M8)} + V_{Dsat(M11)}$ for the P-channel side. If the saturation voltage of the driver FET (M_6) and the cascode FET (M_{12}) are the same, then this requires a bias FET with a current density four times higher. In reality, the body effect in M_{12} and tolerance considerations mean that this factor will have to be somewhat higher than 4 [13].

Hence, there is a requirement for more replicas of the incoming reference current. The whole scheme develops to the configuration shown in Figure 2.74, where the circuit is represented in its simplest ideal form. Note that all FET scaling is applied to device widths; the lengths are the same throughout. In practice, this circuit would have a large offset due to the unequal drain voltages in the various current mirrors, and balancing dummy FETs would be needed.

2.3.6.3 Folded Cascode Amplifiers

A common problem in modern CMOS design is the basing of folded cascode amplifier stages [13]. Folded cascode amplifiers are much used to get reasonable common-mode and output range in low power supply voltage situations. A typical example of a folded cascode amplifier is shown in Figure 2.75.

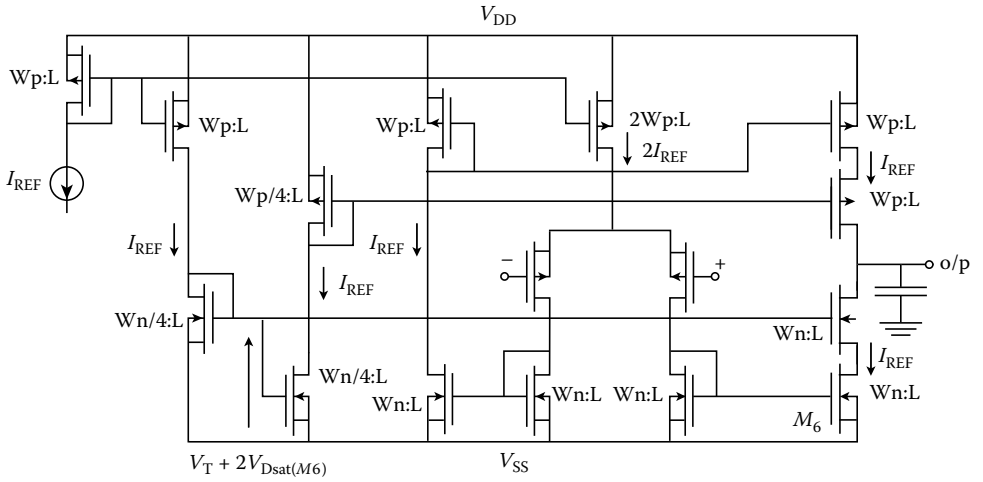


FIGURE 2.74 Cascode simple single-stage amplifier with biasing circuits.

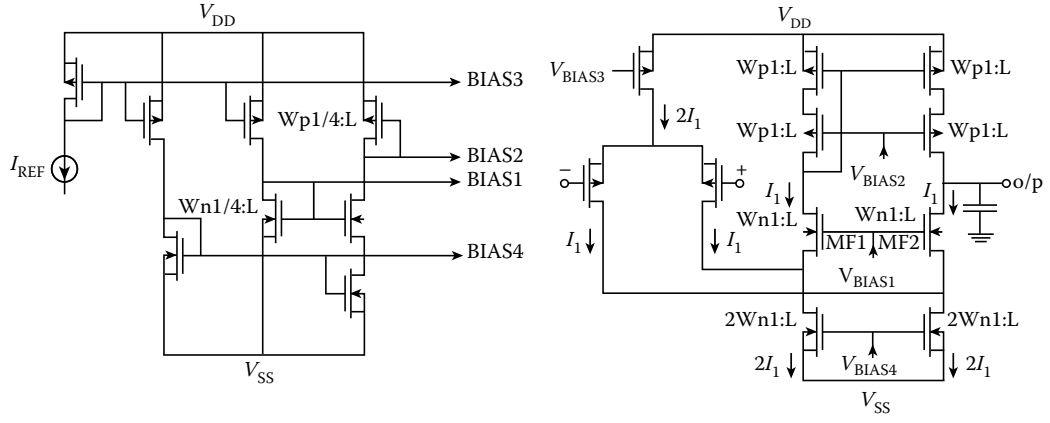


FIGURE 2.75 Folded cascode amplifier with biasing circuits.

The biasing of the folded cascode architecture is basically similar to that of a nonfolded cascode provided that correct current densities are maintained in the FETs. This is important because the folding current source (*MF*1, *MF*2) and the cascode will have different current levels and the bias must allow for this and set the cascode FET to the minimum safe operating bias. A ratio of 1:4 in width for the same current density gives the ideal bias for equal saturation voltages. Differing values of saturation voltage must be summed and the bias FET scaled accordingly.

2.3.6.4 Current Mirrors

The folded cascode amplifier of Figure 2.75 includes the current mirror of Figure 2.76 [18]. This circuit is a high compliance current mirror featuring low input voltage and low minimum output voltage. The cascode devices embedded in it are biased from an FET of width ratio running at the same quiescent current as the mirror. A width ratio of 1:4 is predicted by simple theory for equal saturation voltages in the mirror.

2.3.7 Biasing of Circuits with Low Power Supply Voltage

The topic of low-voltage analog MOS circuits is an important one because of the requirement for battery operation and also for compatibility with advanced digital IC processes with low power supply voltages. In order to maintain a reasonably high dynamic range in low supply voltage analog circuits, it is essential that the circuits operate with signal swings that are a very large fraction of the total supply voltage. Since operational amplifiers are key components in analog circuits, this has led to the design of operational amplifiers with "rail-to-rail" input common-mode voltage and output voltage capability. The design of such an input stage requires new approaches to biasing and we shall give brief details of an example of one such circuit [27].

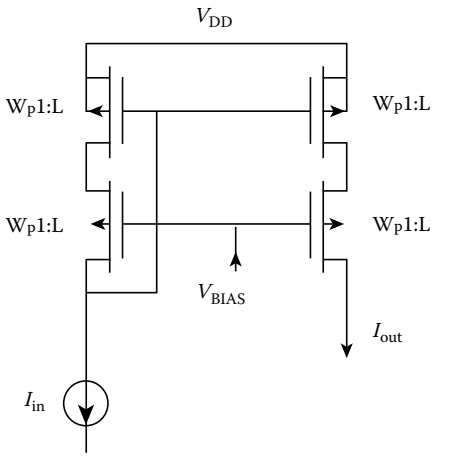


FIGURE 2.76 High-compliance current mirror.

A conventional differential pair of N- or P-channel MOSFETs would not provide sufficient input voltage common-mode range. This is because the input FET pair and the FET realizing the tail current source would tend to come out of saturation at one extreme of input common-mode voltage (negative for NMOS, positive for PMOS). This is overcome by combining an NMOS and PMOS differential pair as shown in Figure 2.77. However, this circuit has the disadvantage that the effective transconductance varies widely with common-mode input voltage since in the middle range, both differential pairs are conducting but at the extremes, only one differential pair is conducting. This produces a common-mode voltage dependence of amplifier dynamic performance and makes it difficult to optimize the dynamic performance for all input conditions.

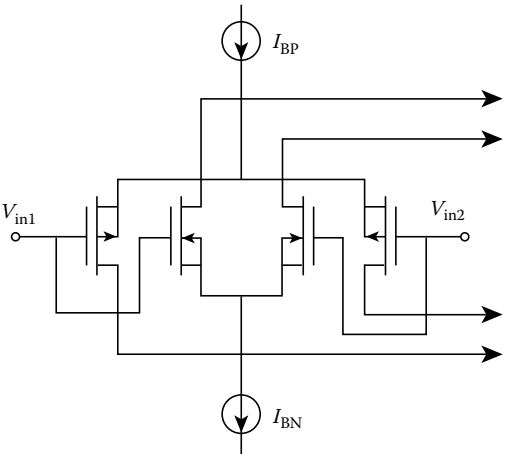


FIGURE 2.77 Op-amp input stage for rail-to-rail operation.

An elegant solution to this problem is reported in Ref. [27]. Assuming that the drain current of a MOSFET can be described by a square-law relationship, then transconductance is proportional to the square root of bias current. Since the overall effective transconductance of the input stage in Figure 2.77 is the sum of the effective transconductance of each pair, it follows that the condition when the overall transconductance is independent of common-mode input voltage is

$$\sqrt{I_{BN}} + \sqrt{I_{BP}} = \text{Constant} \tag{2.127}$$

Bias currents satisfying this relationship can be implemented using the MOS translinear circuit principle [27]. This principle applies to circuits where the gate–source ports of MOSFETs form a closed loop. Assuming that the devices are describable by a square-law drain–current relationship, the sum of the square roots of the drain currents of the MOSFETs whose ports are connected in a clockwise fashion equals the sum of the square roots of the drain currents of the counterclockwise-connected MOSFETs.

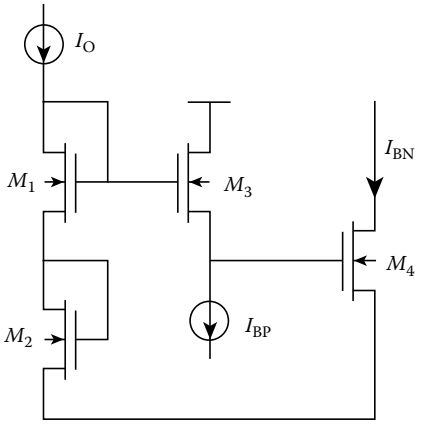


FIGURE 2.78 Bias circuit based on MOS translinear principle.

The application of the basic idea to implement bias currents according to Equation 2.127 is shown in Figure 2.78. Since the clockwise-connected MOSFETs M_1 and M_2 have a constant drain current I_o, the translinear principle implies that the drain current in M_3 and M_4 satisfy Equation 2.127. The development of the schematic bias circuit in Figure 2.78, the input stage in Figure 2.77, and a class *AB* output stage into a fully operational amplifier is described in Ref. [27]. The circuit operates with a minimum power supply of 2.5 V.

2.3.8 Dynamic Biasing

Dynamic biasing is a technique that is applicable to amplifiers in sampled data systems, such as switched capacitor filters and data converters [6,9]. Such amplifiers are required to meet two key requirements. These are fast settling time in order to allow high switching rates and high gain in order to obtain precision performance [13]. Fast settling time is obtained for maximum effective device transconductance and device transconductance g_m is approximately proportional to, $\sqrt{I_B}$, where I_B is bias current. Gain is given by g_m/g_o, where g_o is output conductance. Since g_o is proportional to bias current, gain is "inversely" proportional to $\sqrt{I_B}$. Thus maximum settling time requires a high bias current and maximum gain requires a low bias current. The dynamic bias technique reconciles these two requirements.

Figure 2.79 shows a typical switched capacitor integrator, which is a basic building block for implementing high-order switched capacitor systems. The switches are controlled by two-phase nonoverlapping clock signals ϕ and $\bar{\phi}$. The operational amplifier would generally have a first stage comprising a differential MOSFET pair with constant current source bias. The equivalent of this with dynamic biasing is shown in Figure 2.80, where the constant current source has been replaced by the combination of capacitor *C* and switches S1 and S2. We refer to the integrator of Figure 2.79. During phase ϕ, the capacitor C_1 is being charged up to the input voltage and the amplifier is inactive. Meanwhile, in the dynamically biased amplifier of Figure 2.80, the capacitor *C* is being discharged. In phase $\bar{\phi}$, capacitor C_1 in Figure 2.79 is connected to the input of the amplifier, where the output voltage is required to change to absorb the incoming charge. At the same time, capacitor *C* in Figure 2.80 is connected to the differential pair and immediately starts to conduct a high current. The high current through the amplifier MOSFETs provides a high effective amplifier slew rate and fast initial settling time, although the gain of the amplifier during this initial part of the clock phase is low. As time progresses, capacitor *C* becomes charged and the current in the amplifier MOSFETs reduces. This increases the gain of the amplifier leading to a high precision of the amplifier output voltage.

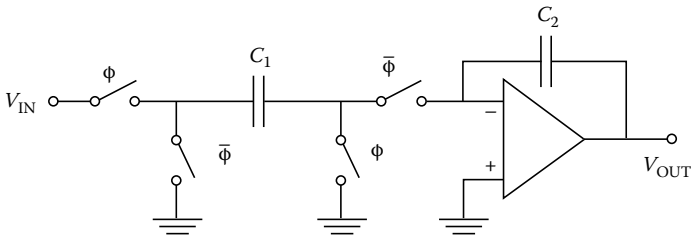


FIGURE 2.79 Typical switched capacitor integrator.

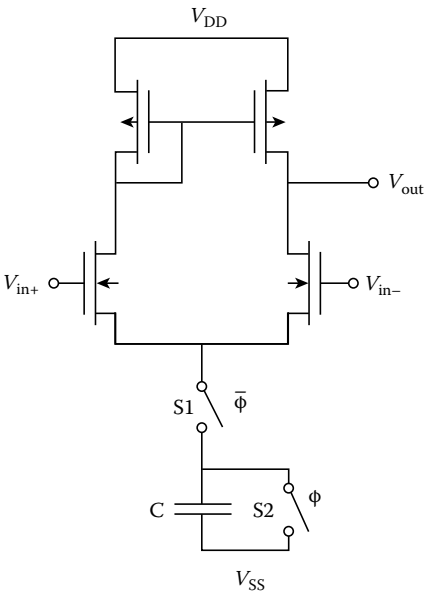


FIGURE 2.80 Simple op-amp with dynamic biasing.

Eventually, the amplifier current falls to zero with the output voltage at this required level. If, as would usually be the case, it is required to sample the amplifier output voltage in both phases, then the dynamic current source comprising capacitor *C* and the two switches S1 and S2 in Figure 2.80 would need to be duplicated with opposite switch phasing. This technique considerably increases the gain available from an amplifier since the effective gain depends on the low bias current condition and is very high. Dynamic biasing may, however, be easily applied to both stages of a two-stage amplifier if required [9]. Also, efficient schemes are available for the dynamic biasing of several amplifiers in a circuit.

Dynamic biasing is well worth considering in sampled date applications, such as switched capacitor filters and data converters. It can maximally exploit a given low power consumption to obtain good dynamic circuit performance. A variant of this approach [11] is adaptive biasing in which the input differential signal is sensed and the bias current is increased for large differential input signals to speed up the slewing response.

2.3.9 Conclusions

The task of designing voltage and current references and bias circuits is an important one. The requirements are very diverse, ranging from high precision, as required in data converters, to moderate, as required in general biasing situations. In these sections, space has been sufficient only to discuss some outstanding work in the area and some of the main principles. It is hoped that the reader will consult the references for more detailed information.

References

1. S. M. Sze, *Physics of Semiconductor Devices*, New York: Wiley Interscience, 1969.
2. R. J. Widlar, New developments in IC voltage regulators, *IEEE J. Solid-State Circuits*, SC-6, 2–7, Feb. 1971.
3. Y. P. Tsividis, A CMOS voltage reference, *IEEE J. Solid-State Circuits*, SC-13, 774–778, Dec. 1978.
4. R. A. Blauschild, P. A. Tucci, R. S. Muller, and R. G. Meyer, A new NMOS temperature-stable voltage reference, *IEEE J. Solid-State Circuits*, SC-13, 767–773, Dec. 1978.
5. M. E. Hoff, J. Huggins, and B. M. Warren, An NMOS telephone Codec for transmission and switching applications, *IEEE J. Solid-State Circuits*, SC-14, 47–50, Feb. 1979.
6. M. A. Copelend and J. M. Rabaey, Dynamic amplifier for MOS technology, *Electron. Lett.*, 15, 301, 302, May 1979.
7. E. A. Vittoz and O. Neyroud, A low voltage CMOS bandgap reference, *IEEE J. Solid-State Circuits*, SC-14, 573–577, Jun. 1979.
8. H. I. Oguey and B. Gerber, MOS voltage reference based on polysilicon gate work function difference, *IEEE J. Solid-State Circuits*, SC-15, 264–269, Jun. 1980.
9. B. J. Hosticka, Dynamic CMOS amplifiers, *IEEE J. Solid-State Circuits*, SC-15, 887–894, Oct. 1980.
10. R. Ye and T. Tsividis, Bandgap voltage reference sources in CMOS technology, *Electron. Lett.*, 18(1), 24, 25, Jan. 7, 1982.

11. M. C. Degruwe, J. Rijmenants, E. A. Vittoz, and H. J. de Man, Adaptive biasing CMOS amplifiers, *IEEE J. Solid-State Circuits*, SC-17, 552–528, Oct. 1980.
12. B.-S. Song and P. R. Gray, A precision curvature-compensated CMOS bandgap reference, *IEEE J. Solid-State Circuits*, SC-18, 634–643, Dec. 1983.
13. P. R. Gray and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, New York: Wiley, 1984, pp. 730–737.
14. J. Michejda and S. K. Kim, A precision CMOS bandgap reference, *IEEE J. Solid-State Circuits*, SC-19, 1014–1021, Dec. 1984.
15. B. J. Hosticka, K.-G. Dalsab, D. Krey, and G. Zimmer, Behavior of analog MOS integrated circuits at high temperatures, *IEEE J. Solid-State Circuits*, SC-20, 871–874, Aug. 1985.
16. M. G. K. R. Degrauwe, O. N. Leuthold, E. A. Vittoz, H. J. Oguey, and A. Descombes, CMOS voltage reference using lateral bipolar transistors, *IEEE J. Solid-State Circuits*, SC-20, 1151–1157, Dec. 1985.
17. S. L. Lin and C. A. T. Salama, A $V_{be}(T)$ model with application to bandgap reference design, *IEEE J. Solid-State Circuits*, SC-20, 1283–1285, Dec. 1985.
18. A. J. J. Boudewijns, Amplifier arrangement, U.S. Patent 4,893,090, Granted Jan. 8, 1990 (submitted Sept. 1988).
19. W. M. Sansen, F. O. Eynde, and M. Steyaert, A CMOS temperature-compensated current reference, *IEEE J. Solid-State Circuits*, SC-23, 821–824, Jun. 1988.
20. M. Ferro, F. Salerno, and R. Castello, A floating CMOS bandgap voltage reference for differential applications, *IEEE J. Solid-State Circuits*, SC-24, 690–697, Jun. 1989.
21. J. M. Steininger, Understanding wideband MOS transistors, *IEEE Circuits Devices Mag.*, 6, 26–31, May 1990.
22. G. Nicollini and D. Senderowicz, A CMOS bandgap reference for differential signal processing, *IEEE J. Solid-State Circuits*, SC-21, 41–50, Jan. 1991.
23. K. Ishibashi and K. Sasaki, A voltage down converter with submicroampere standby current for low power static RAMs, *IEEE J. Solid-State Circuits*, SC-27, 920–925, June 1992.
24. C.-Y. Wu and S.-Y. Chin, High precision curvature-compensated CMOS band-gap voltage and current references, *J. Analog Integrat. Circuits Signal Process.*, 2(3), 207–215, Sept. 1992.
25. S. D. Willingham and K. W. Martin, A BiCMOS low distortion 8 MHz low-pass filter, *IEEE J. Solid-State Circuits*, SC-28, 1234–1245, Dec. 1993.
26. J. Choma, Jr., Temperature stable voltage controlled current source, *IEEE Trans. Circuits Syst. I*, 41, 405–411, May 1994.
27. J. H. Botma, R. Jiegerink, S. L. J. Gierkink, and R. F. Wassenaar, Rail-to-rail constant Gm input stage and class AB output stage for low-voltage CMOS op amps, *Analog Integrat. Circuits Signal Process.*, 6(2), 121–133, Sept. 1994.

2.4 Canonical Cells of MOSFET Technology

Mohammed Ismail, Shu-Chuan Huang, Chung-Chih Hung, and Trond Saether

Analog integrated circuits have long been designed in technologies other than CMOS. But modern analog and mixed-signal VLSI applications in areas such as telecommunications, smart sensors, battery-operated consumer electronics, and artificial neural computation require CMOS analog design solutions. In recent years, analog CMOS circuit design has shown signs of dramatic change. Field programmable analog arrays and modular analog VLSI circuits [1] are representatives of emerging analog design philosophies leading to a whole new generation of analog circuit and layout design methodologies.

This section discusses basic cells used in contemporary CMOS analog integrated circuits. The performance of a CMOS (bipolar) circuit can often be improved further by incorporating a limited

number of bipolar (CMOS) transistors on the same substrate. The resulting circuits are called BiCMOS circuits. BiCMOS circuits that are predominantly CMOS will also be discussed. First, we discuss primitive analog cells. These cells may or may not require device matching for proper operation. Second, we introduce modern and simple circuit techniques to mitigate nonideal effects and significantly improve circuit performance, and finally, we discuss basic voltage amplifier circuits. The presented cells will help in the systematic design of analog integrated circuits and could constitute an efficient analog VLSI cell library. Throughout this section, MOS transistors are assumed to be biased in strong inversion.

2.4.1 Matched Device Pairs

Figure 2.81 shows basic MOS transistors pairs [2] operating in the saturation region, where only NMOS transistors pairs are shown. Figure 2.81a shows a differential pair with no direct connection between the two transistors. The resultant differential pair is characterized by the difference in the drain currents (using the simple square-law equation); that is

$$\begin{aligned} I_{a1} - I_{a2} &= \frac{K}{2}(V_{G1} - V_{S1} - V_T)^2 - \frac{K}{2}(V_{G2} - V_{S2} - V_T)^2 \\ &= \frac{K}{2}[(V_{G1} - V_{G2}) - (V_{S1} - V_{S2})] \\ &\quad \times [(V_{G1} + V_{G2}) - (V_{S1} + V_{S2}) - 2V_T] \end{aligned} \tag{2.128}$$

where $K(=\mu C_{ox}W/L)$ and V_T are the transconductor parameter and the threshold voltage of the transistor, respectively. Figure 2.81b is a common-source or source-coupled differential pair, a special case of circuit (a) with $V_{S1} = V_{S2} = V_S$, and the differential current is

$$I_{b1} - I_{b2} = \frac{K}{2}(V_{G1} - V_{G2})[(V_{G1} + V_{G2}) - 2V_S - 2V_T] \tag{2.129}$$

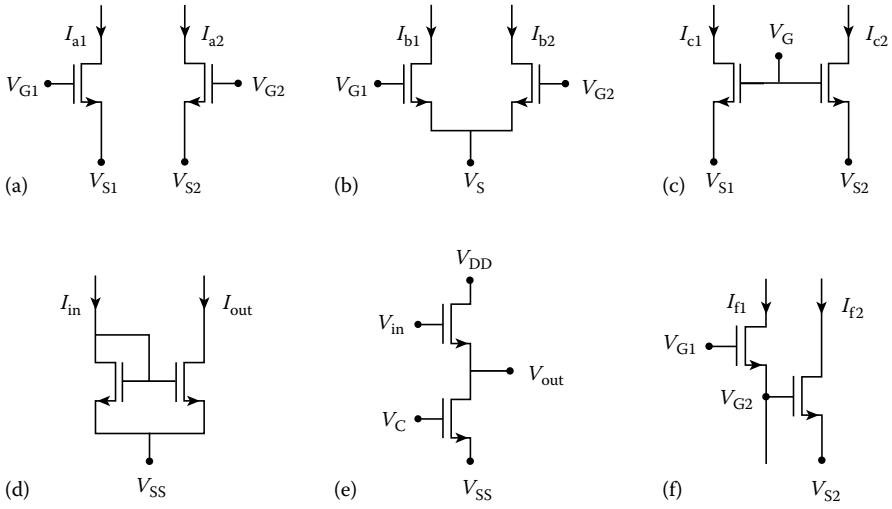


FIGURE 2.81 Matched primitive cells operating in the saturation region. (a) A differential pair with no direct connection between the two transistors, (b) is a common source or source coupled differential pair, (c) a common-gate differential pair, (d) a well-known simple current mirror, (e) a voltage follower, and (f) a rearranged transistor pair.

Figure 2.81c is a common-gate differential pair with $V_{G1} = V_{G2} = V_G$ in circuit (a), and the differential current is obtained as

$$I_{c1} - I_{c2} = -\frac{K}{2}(V_{S1} - V_{S2})[2V_G - (V_{S1} + V_{S2}) - 2V_T] \tag{2.130}$$

Differential pairs are essential building blocks of circuits such as op-amps, differential difference amplifiers and operational transconductance amplifiers. Several linear *V–I* converters built by these cells have been developed.

Current mirrors are usually used as loads for amplifier stages. Moreover, current mirrors are essential building blocks in modern current-mode analog integrated circuits. Figure 2.81d shows a well-known simple current mirror. Ideally, the input current I_{in} is equal to the output current I_{out} for matched transistors. In practice, a nonunity I_{out} to I_{in} ratio occurs due to finite output resistance resulting from channel length modulation effects. The output resistance can be increased by Wilson or cascode current mirrors at the expense of a limited output swing, which is not desired in low-voltage applications. A regulated current mirror can improve both the output resistance and swing but increase circuit complexity. Detail analysis and comparison are discussed in Ref. [3].

A voltage follower is shown in Figure 2.81e. Since the same current flows in both transistors, their gate–source voltages are the same. That is,

$$V_{in} - V_{out} = V_C - V_{SS} \tag{2.131}$$

and therefore

$$V_{out} = V_{in} - V_C + V_{SS} \tag{2.132}$$

Alternatively, a transistor pair can be arranged as the circuit shown in Figure 2.81f, which is used as a basic cell for composite MOSFET (COMFET) circuits [4]. The differential current is given by

$$\begin{aligned} I_{f1} - I_{f2} &= \frac{K}{2}(V_{G1} - V_{G2} - V_T)^2 - \frac{K}{2}(V_{G2} - V_{S2} - V_T)^2 \\ &= \frac{K}{2}(V_{G1} - V_{S2} - 2V_T)(V_{G1} - 2V_{G2} + V_{S2}) \end{aligned} \tag{2.133}$$

With proper biasing, linear *V–I* conversion can be achieved by this transistor cell.

Transistor pairs operating in the triode region are found mostly in simulating linear transconductors and resistors, for example, those in MOSFET-C filters. Figure 2.82 shows three popular examples, where the nonlinear terms in the drain current equations are canceled. A simple drain current equation in the triode region is

$$I_D = K\left[(V_G - V_T)(V_D - V_S) - \frac{1}{2}\left(V_D^2 - V_S^2\right)\right] \tag{2.134}$$

$$= \frac{K}{2}(V_G - V_S - V_T)^2 - \frac{K}{2}(V_G - V_D - V_T)^2 \tag{2.135}$$

Equation 2.135 gives another form of the triode current equation. In some cases, circuit analysis can be performed more easily with this form than using Equation 2.134. The resulting current equations of the equivalent "MOS resistors" are now obtained.

For the circuit shown in Figure 2.82a, a two-transistor transconductor, the current difference is given by [5]

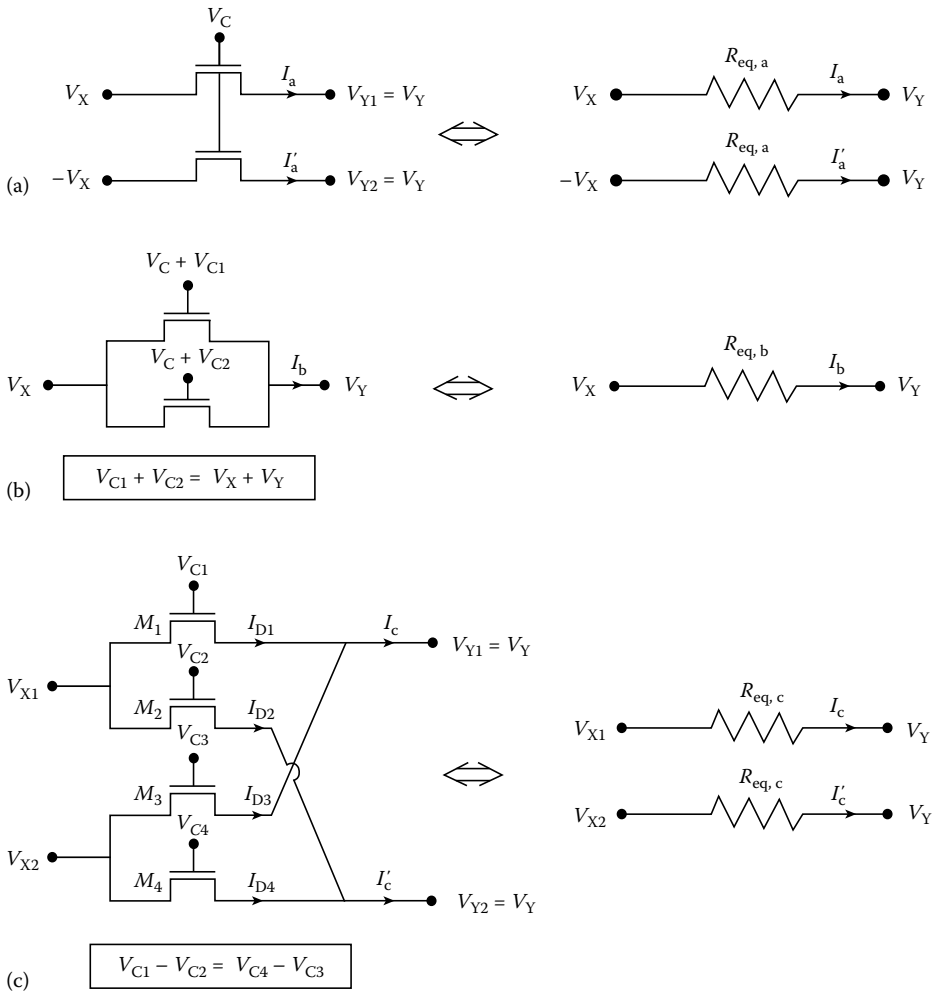


FIGURE 2.82 Matched primitive cells operating in the triode region. Three popular examples: (a) a two-transistor transconductor, (b) a realized floating resistor, and (c) a four-transistor transconductor.

$$\begin{aligned} I_a - I'_a &= K\left[(V_C - V_T)(V_X - V_Y) - \frac{1}{2}\left(V_X^2 - V_Y^2\right)\right] \\ &\quad - K\left[(V_C - V_T)(-V_X - V_Y) - \frac{1}{2}\left(V_X^2 - V_Y^2\right)\right] \\ &= 2K(V_C - V_T)V_X \end{aligned} \tag{2.136}$$

and the equivalent resistance is obtained as

$$R_{eq,a} = \frac{2V_X}{I_a - I'_a} = \frac{1}{K(V_C - V_T)} \tag{2.137}$$

The circuit shown in Figure 2.82b realizes a floating resistor [5], where

$$
\begin{aligned}
I_b &= K\left[(V_C + V_{C1} - V_T)(V_X - V_Y) - \frac{1}{2}\left(V_X^2 - V_Y^2\right)\right] \\
&\quad + K\left[(V_C + V_{C2} - V_T)(V_X - V_Y) - \frac{1}{2}\left(V_X^2 - V_Y^2\right)\right] \\
&= K\left[(2V_C - 2V_T + V_X + V_Y)(V_X - V_Y) - \left(V_X^2 - V_Y^2\right)\right] \\
&= 2K(V_C - V_T)(V_X - V_Y)
\end{aligned} \tag{2.138}
$$

and the equivalent resistance is

$$
R_{eq,b} = \frac{V_X - V_Y}{I_b} = \frac{1}{2K(V_C - V_T)} \tag{2.139}
$$

An implementation with $V_{C1} = V_{C2} = (V_X + V_Y)/2$ has been described in Ref. [6].

The circuit shown in Figure 2.82c, a four-transistor transconductor [7], gives the following current equation:

$$
\begin{aligned}
I_c - I_c' &= K\left[(V_{C1} - V_T)(V_{X1} - V_Y) - \frac{1}{2}\left(V_{X1}^2 - V_Y^2\right)\right] \\
&\quad - K\left[(V_{C2} - V_T)(V_{X1} - V_Y) - \frac{1}{2}\left(V_{X1}^2 - V_Y^2\right)\right] \\
&\quad + K\left[(V_{C3} - V_T)(V_{X2} - V_Y) - \frac{1}{2}\left(V_{X2}^2 - V_Y^2\right)\right] \\
&\quad - K\left[(V_{C4} - V_T)(V_{X2} - V_Y) - \frac{1}{2}\left(V_{X2}^2 - V_Y^2\right)\right]
\end{aligned} \tag{2.140}
$$

$$
\begin{aligned}
&= K(V_{C1} - V_{C2})(V_{X1} - V_{X2}) \\
&= K(V_{C4} - V_{C3})(V_{X1} - V_{X2})
\end{aligned} \tag{2.141}
$$

and

$$
R_{eq,c} = \frac{V_{X1} - V_{X2}}{I_c - I_c'} = \frac{1}{K(V_{C1} - V_{C2})} = \frac{1}{K(V_{C4} - V_{C3})} \tag{2.142}
$$

Note that I_c and I_c' are taken at the V_Y nodes. It is very interesting to know that nonlinearity cancellation is also achieved with the four transistors operating in the saturation region [2].

Figure 2.82a and c are usually used together with op-amps to simulate resistors, where the virtual short property of op-amps makes $V_{Y1} = V_{Y2}$.

2.4.2 Unmatched Device Pairs

Figure 2.83 shows primitive cells that do not require matching, unless specified. Figure 2.83a and b are parallel and series composite NMOS transistors, respectively, which are very useful in laying out very wide or long transistors, respectively. The equivalent device transconductance parameter K_{eq} is calculated as follows:

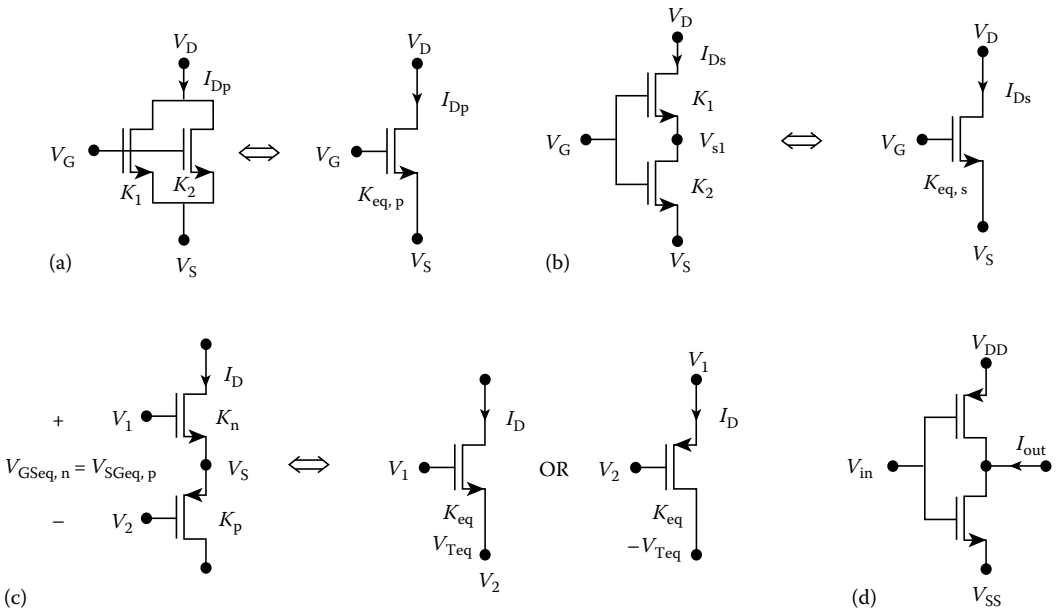


FIGURE 2.83 Unmatched primitive cells. (a) A parallel composite NMOS transistor, (b) a series composite NMOS transistor, (c) a CMOS composite transistor, and (d) a CMOS inverter.

For the parallel composite transistor, the drain current is written as

$$\begin{aligned} I_{Dp} &= \frac{K_{eq,p}}{2}(V_G - V_S - V_T)^2 \\ &= \frac{K_1}{2}(V_G - V_S - V_T)^2 + \frac{K_2}{2}(V_G - V_S - V_T)^2 \end{aligned} \tag{2.143}$$

That is,

$$K_{eq,p} = K_1 + K_2 \tag{2.144}$$

or alternatively

$$\left(\frac{W}{L}\right)_{eq,p} = \left(\frac{W}{L}\right)_1 + \left(\frac{W}{L}\right)_2 \tag{2.145}$$

Using the same channel length L, it can be simplified as

$$W_{eq,p} = W_1 + W_2 \tag{2.146}$$

As a result, a wider transistor can be realized by parallel connection of two or more narrower transistors.

For the series composite transistor, note that the lower transistor is always operating in the triode region due to the requirement $V_G - V_{S1} > V_T$, to turn on the upper transistor. The resultant drain current is given by

$$I_{Ds} = \frac{K_{eq,s}}{2}(V_G - V_S - V_T)^2 \tag{2.147}$$

$$= \frac{K_1}{2}(V_G - V_{S1} - V_T)^2 \tag{2.148}$$

$$= \frac{K_2}{2}\left[(V_G - V_S - V_T)^2 - (V_G - V_{S1} - V_T)^2\right] \tag{2.149}$$

From the preceding equations, we have

$$(V_G - V_S - V_T)^2 = \frac{2I_{Ds}}{K_{eq,s}} \tag{2.150}$$

$$(V_G - V_{S1} - V_T)^2 = \frac{2I_{Ds}}{K_1} \tag{2.151}$$

Substituting the preceding equations into Equation 2.149, we obtain

$$I_{Ds} = \frac{K_2}{2}\left(\frac{2I_{Ds}}{K_{eq,s}} - \frac{2I_{Ds}}{K_1}\right) \tag{2.152}$$

That is,

$$\frac{1}{K_{eq,s}} = \frac{1}{K_1} + \frac{1}{K_2} \tag{2.153}$$

or

$$\left(\frac{L}{W}\right)_{eq,s} = \left(\frac{L}{W}\right)_1 + \left(\frac{L}{W}\right)_2 \tag{2.154}$$

Similarly, with fixed channel width W, the above equation is simply obtained as

$$L_{eq,s} = L_1 + L_2 \tag{2.155}$$

which indicates that the equivalent transistor can be used to realize a long-channel device with shorter channel ones.

Figure 2.83c is a CMOS composite transistor, which can be seen as equivalent to either an NMOS or a PMOS transistor operating in the saturation region. In contrast, the composite transistors shown in Figure 2.83a and b can operate in both saturation and triode regions. The main advantage of the equivalent composite transistor shown in Figure 2.83c is that both their equivalent gate and source nodes have high input impedances, which is desired in some circuits. The equivalent K and V_T are obtained by the equations of the gate–source voltages.

$$\begin{aligned} V_1 - V_2 &= V_{GSn} + V_{SGp} \\ &= \sqrt{\frac{2I_D}{K_n}} + \sqrt{\frac{2I_D}{K_p}} + V_{Tn} - V_{Tp} \\ &= \sqrt{\frac{2I_D}{K_{eq}}} + V_{Teq} \end{aligned} \tag{2.156}$$

which give

$$\frac{1}{\sqrt{K_{eq}}} = \frac{1}{\sqrt{K_n}} + \frac{1}{\sqrt{K_p}} \tag{2.157}$$

and

$$V_{Teq} = V_{Tn} - V_{Tp} \tag{2.158}$$

Finally, Figure 2.83d shows a CMOS inverter, which could be used as a transconductor [8]. Its output current is

$$\begin{aligned} I_{out} &= \frac{K_n}{2}(V_{in} - V_{SS} - V_{Tn})^2 - \frac{K_p}{2}(V_{DD} - V_{in} + V_{Tp})^2 \\ &= a(V_{in} - V_{Tn})^2 + bV_{in} + c \end{aligned} \tag{2.159}$$

where

$$\begin{aligned} a &= \frac{1}{2}(K_n - K_p) \\ b &= -K_n V_{SS} + K_P(V_{DD} - V_{Tn} + V_{Tp}) \\ c &= \frac{K_n}{2}\left(2V_{SS}V_{Tn} + V_{SS}^2\right) + \frac{K_p}{2}\left[V_{Tn}^2 - (V_{DD} + V_{Tp})^2\right] \end{aligned}$$

2.4.3 Composite Transistors

The body effect of a transistor is due to nonzero source to bulk voltage (V_{SB}), which widens the depletion region between the source and bulk and therefore increases the absolute value of its threshold voltage. The threshold voltage (referred to the source) is dependent on V_{SB} and is given by

$$\begin{aligned} V_{Tn} &= V_{Tno} + \gamma\left(\sqrt{2|\phi_F| + V_{SB}} - \sqrt{2|\phi_F|}\right) \quad \text{for NMOS} \\ V_{Tp} &= V_{Tpo} - \gamma\left(\sqrt{2|\phi_F| + V_{SB}} - \sqrt{2|\phi_F|}\right) \quad \text{for PMOS} \end{aligned}$$

where $2|\phi_F|$ is the potential required for strong inversion and γ is the body effect parameter.

Usually, bulk regions of an NMOS transistor and a PMOS transistor are tied to the most negative voltage (V_{SS}) and the most positive voltage (V_{DD}) respectively to turn off the parasitic diodes associated with source-bulk and drain-bulk. In some cases, bulk regions are directly connected to transistor sources ($V_{SB} = 0$) to eliminate the body effect; for example, in the follower of Figure 2.81e, the bulk must be connected to the source in each transistor to ensure equal threshold voltages. This is achieved by putting each device in a separated well, which must be the P-well for NMOS devices. However, separate wells require large layout areas. Besides, unless twin-tub processes are used, only one type of transistor (either NMOS or PMOS depending on the process) can be connected this way.

Due to the body effect, the equivalent threshold voltage of a CMOS composite transistor would be large (two threshold voltages plus extra voltage resulting from the body effect), which would render it unsuitable for low-voltage applications. The equivalent threshold voltage could be reduced by replacing one of the MOS transistors with a BJT, as shown in Figure 2.84 [9]. For the stacked composite BiCMOS transistors, the equivalent threshold voltage is given by $V_{Teq} \approx |V_T| + 0.7$ V, where V_T is the threshold

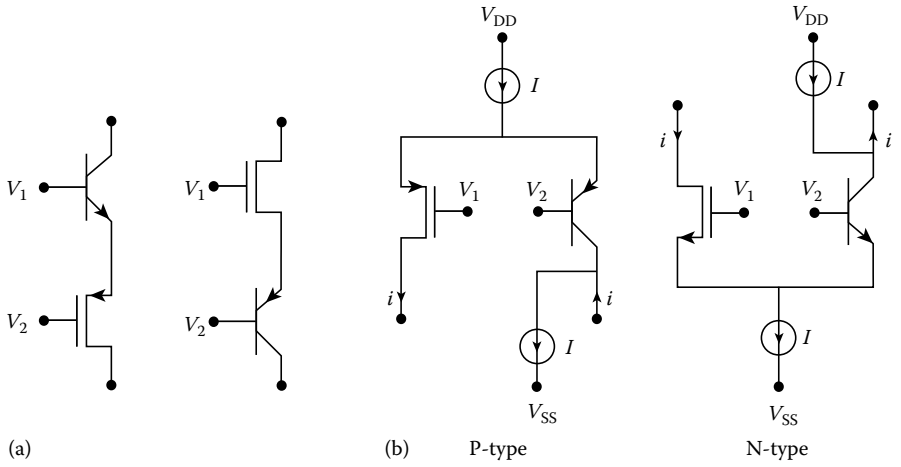


FIGURE 2.84 BiCMOS composite transistors: (a) stacked version and (b) folded version.

voltage of the NMOS or PMOS transistor, and 0.7 V is the BJT turn-on voltage V_{BE}, which is not subject to body effects. It can be further reduced by the folded arrangement as shown in Figure 2.84b, where $V_{Teq} \approx |V_T| - 0.7$ V. An all-MOS-folded composite transistor can be implemented in a similar manner as shown in Figure 2.85 [10], where $K_2 \gg K_1$. As a result,

$$K_{eq} \approx K_1$$

$$V_{Teq} = V_T - V_{GS2} \tag{2.160}$$

$$\approx -\sqrt{\frac{2I}{K_2}} \tag{2.161}$$

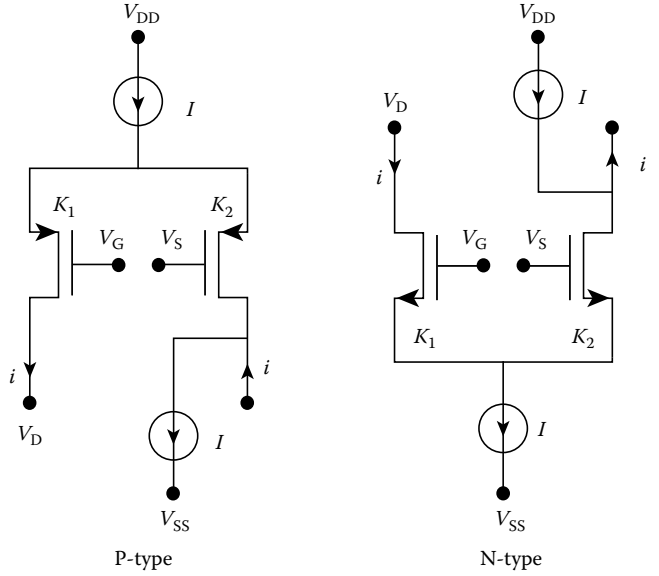


FIGURE 2.85 MOS-folded composite transistors.

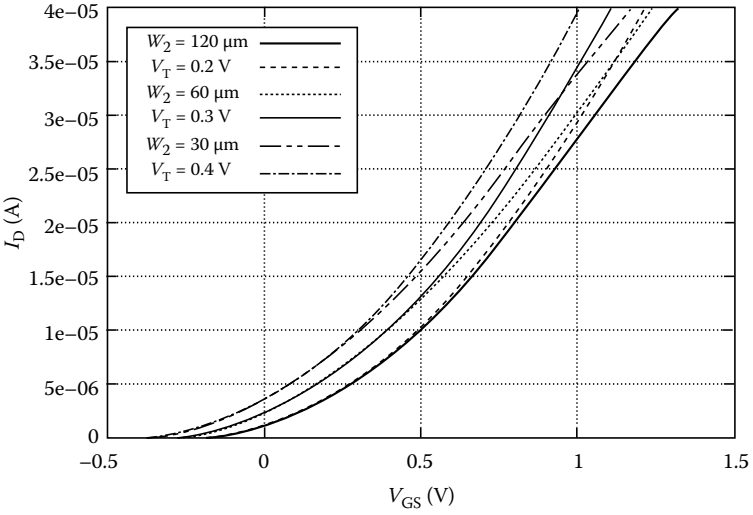


FIGURE 2.86 I_D curves for a folded N-type transistor with various W_2 and a single transistor with various V_T.

where $V_{GS2} \approx \sqrt{(2I)/K_2} + V_T$. Simulation program with integrated circuit emphasis (SPICE) simulation results for the N-type folded transistors ($W_1 = L_1 = L_2 = 3$ μm) operating in the saturation region ($V_{DG} = 0$) with various W_2 are compared with a single NMOS depletion transistor ($W = L = 3$ μm) with various V_T shown in Figure 2.86. It can be seen that a smaller K_2 results in a smaller V_{Teq} (more negative), but with a larger K_2 the composite transistor behaves more like a single transistor having a smaller $|V_{Teq}|$, which could be useful in low-voltage applications.

Forcing $V_{SB} = 0$ to eliminate the body effect is usable only for MOS circuits conducting currents in a single direction. For the circuits shown in Figure 2.82, since resistors operate bidirectionally, the bulk regions of each transistor must be connected to the rail to assure that parasitic diodes are turned off when currents flow in either direction. In fact, the transistors are operating in the triode region symmetrically between drain and source and biased at $V_{DS} \approx 0$. It would, however, result in nonzero V_{SB} and increase the threshold voltage, which increases the equivalent resistance, in Figure 2.82a and b, but introduces nonlinearities. To overcome this problem, one may configure two transistors into one composite transistor as shown in Figure 2.87a, where the bulks of the transistors are interconnected to node V_{S1}. Due to symmetry, this composite transistor can be operated in either direction. Its physical cross section is shown in Figure 2.87b, where the diodes represent the p–n junctions composed by bulk and source/drain nodes. This configuration is equivalent to Figure 2.87c, and one can find that the parasitic diode connected between V_{S1} and V_S would turn on when $V_{S1} - V_S$ is larger than the turn-on voltage of the parasitic diode. However, this is undesired, but fortunately the diode current is restricted by the drain current of M_2. This effect can be illustrated more clearly through the comparison of transistors with various bulk connections, as shown in Figure 2.88, where Figure 2.88d and e are composite transistors (same as Figures 2.87 and Figure 2.83b, respectively), which simulate single transistors. With $V_{DG} = 0$ (diode connection), $V_S = 0$ and $V_{SS} = -5$ V, Figure 2.89 gives simulation results of drain currents for the circuits, shown in (a) and (c)–(e), where the transistor sizes for the circuits, shown in (a)–(c), (d), and (e) are 20/3 μm, 30/3 μm, and 36/3 μm, respectively. One can observe that the curve (c) shown in Figure 2.89 completely departs from curve (a), due to the body effect. Curve (e) fits perfectly to curve (a). Although behaving slightly differently from curve (a), circuit (d) approximates a single transistor as well.

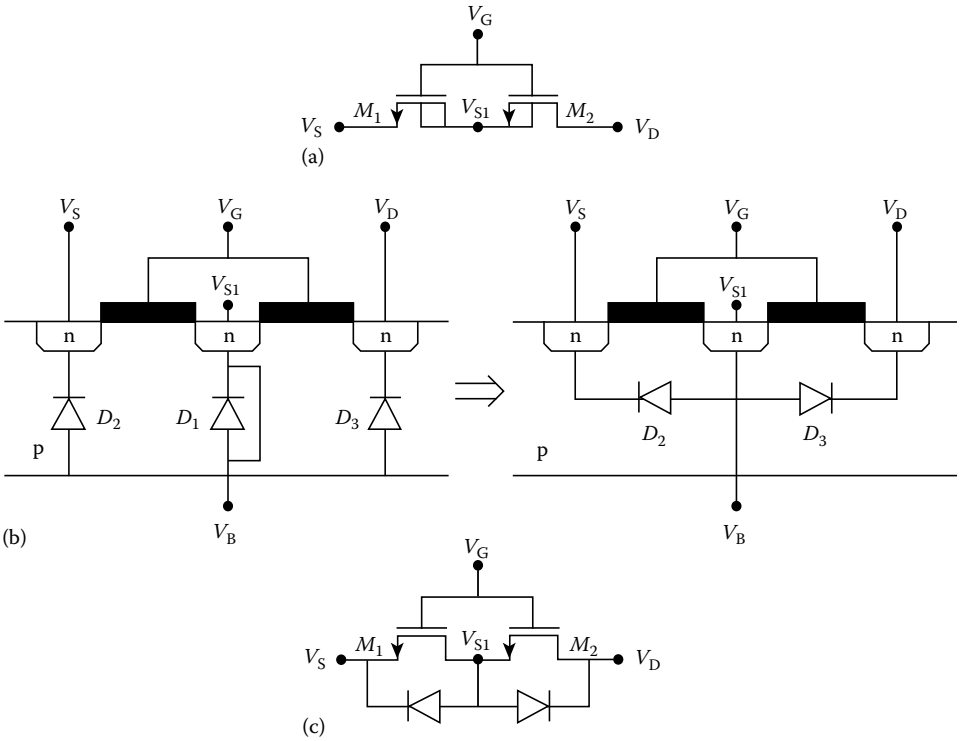


FIGURE 2.87 (a) Composite bidirectional transistor with reduced body effect. (b) Cross-sectional view of the physical device, where the short connection across D_1 is, in effect, placing D_2 and D_3 back-to-back between V_D and V_S and forming a parasitic symmetrical bipolar device. (c) Equivalent circuit of (a). (From Huang, S.-C., Systematic design solutions for analog VLSI circuits, PhD dissertation, Department of Electrical Engineering, Ohio State University, Columbus, OH, 1994. With permission.)

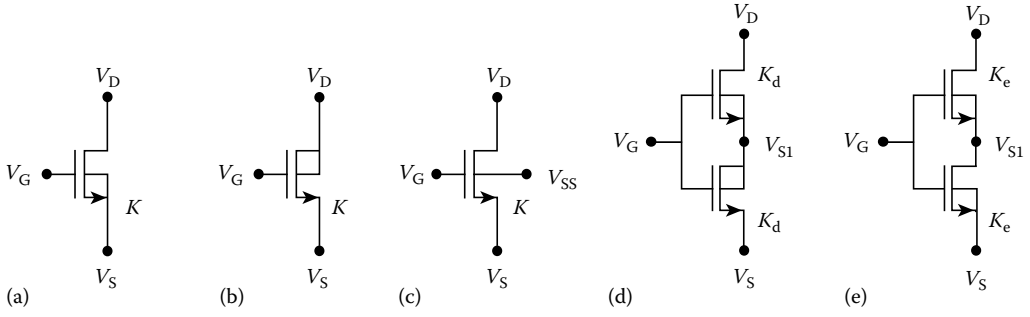


FIGURE 2.88 Transistors with various bulk connections.

Figure 2.90 shows the drain current for the circuit labeled (b), whose current is much larger than those of the rest of the circuits shown in Figure 2.88. Since the bulk of the circuit shown in Figure 2.88b is connected to its drain and the parasitic diode between the drain and bulk is on, the current is dominated by the diode current due to its exponential nature. By using it as in Figure 2.88d, the diode current is limited to the current level of an MOS transistor. This can be seen from Figure 2.91, showing V_{S1} of the

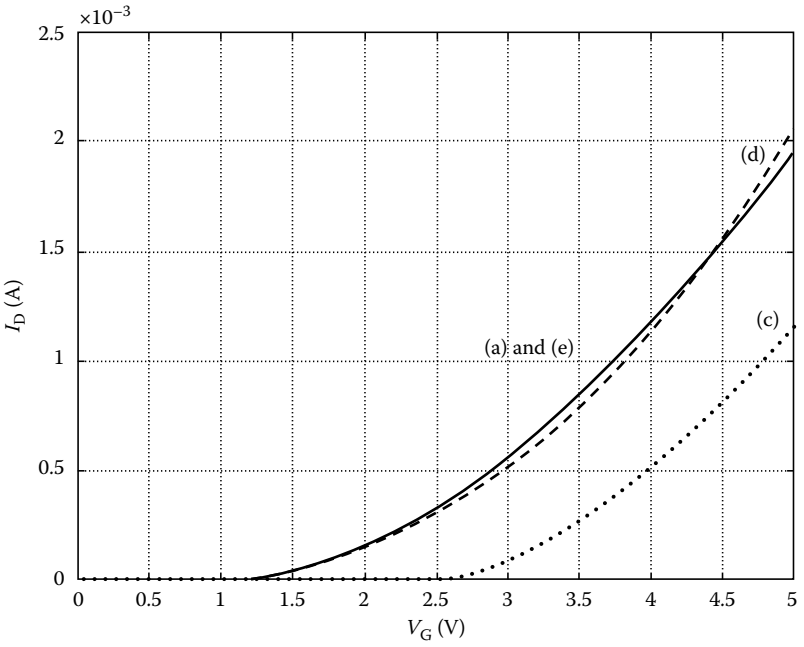


FIGURE 2.89 I_D curves for transistors with various bulk connections.

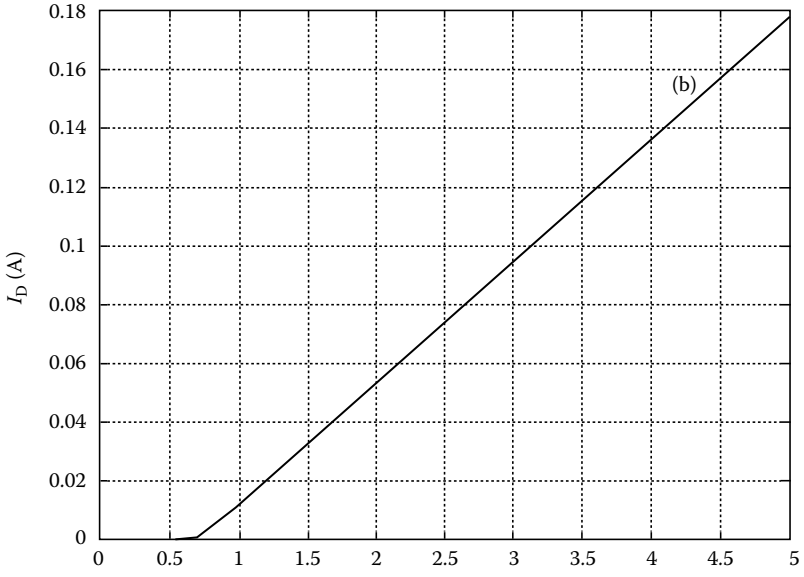


FIGURE 2.90 I_D curve for transistor with bulk connected to drain.

circuit (d) saturated at a voltage ∼0.7 V, which is close to the turn-on voltage of a p–n junction diode. The transistor sizes of the circuits labeled (d) and (e) are adjusted to achieve the same K value of the single transistor (a). According to Equation 2.153, $1/K_{eq,s} = 1/K_1 + 1/K_2$. That is, to achieve $K_{eq,s} = K$, K_e $(= K_1 = K_2)$ is given by

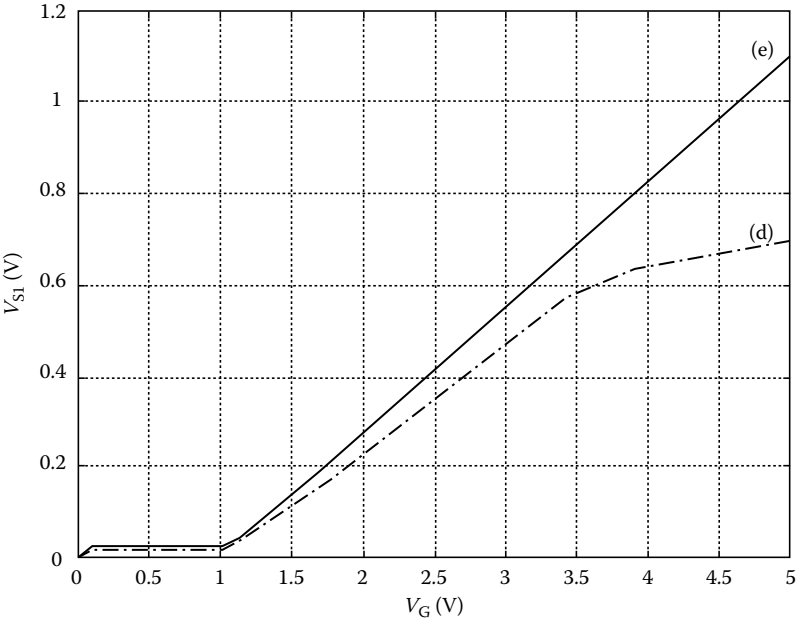


FIGURE 2.91 V_{S1} curves for composite transistors, labeled (d) and (e).

$$K_e = 2K \tag{2.162}$$

For circuit (d), K_d is obtained by rewriting Equation 2.149, which follows that

$$\begin{aligned} I_{Ds} &= \frac{K}{2}(V_G - V_S - V_T)^2 \\ &= \frac{K_d}{2}(V_G - V_{S1} - V_T)^2 \\ &= \frac{K_d}{2}\left[(V_G - V_S - V_T)^2 - (V_G - V_{S1} - V_T)^2\right] + I_s e^{(V_{S1}-V_S)/U_T} \end{aligned}$$

where I_s is the leakage current of the diode and U_T is the thermal voltage. Therefore,

$$\begin{aligned} I_{Ds} &= \frac{K_d}{2}\left(\frac{2I_{Ds}}{K} - \frac{2I_{Ds}}{K_d}\right) + I_s e^{(V_{S1}-V_S)/U_T} \\ &= I_{Ds}\left(\frac{K_d}{K} - 1\right) + I_s e^{(V_{S1}-V_S)/U_T} \end{aligned} \tag{2.163}$$

Divided by I_{Ds}, the preceding equation becomes

$$\begin{aligned} 1 &= \frac{K_d}{K} - 1 + \frac{I_s}{I_{Ds}} e^{(V_{S1}-V_S)/U_T} \\ 2K &= K_d + K\frac{I_s}{I_{Ds}} e^{(V_{S1}-V_S)/U_T} \end{aligned} \tag{2.164}$$

and hence

$$K_d < 2K \tag{2.165}$$

due to the parasitic diode. A SPICE level 2 model is used in the simulation and its higher order effects result in using a device size of 36/3 μm($K_e \neq 2K$), instead of 40/3 μm.

2.4.4 Super MOS Transistors

The channel length modulation effect models the channel shortening effect in the saturation region due to the increase in the depletion width near the drain when increasing V_{DS}. It is modeled as

$$I_D = \frac{K}{2}(V_{GS} - V_T)^2(1 + \lambda V_{DS}) \tag{2.166}$$

where λ is the channel length modulation parameter. The effect results in a finite output impedance of a transistor, since the output impedance is given by

$$r_o = \left(\frac{\partial I_D}{\partial V_{DS}}\right)^{-1} = \frac{1}{\lambda \frac{K}{2}(V_{GS} - V_T)^2} \simeq \frac{1}{\lambda I_D} \tag{2.167}$$

As mentioned previously, this effect would cause inaccuracy in the single current mirror shown in Figure 2.81d, and can be mitigated by using cascode, improved Wilson, or regulated current mirrors as shown in Figure 2.92 [3]. These are based on the gain-boosting principle as shown in Figure 2.93 [11], where for the cascode stage in (a) (used in the cascode and the Wilson current mirrors) the output impedance is given by

$$r_{o,a} = (g_{m2}r_{o2} + 1)r_{o1} + r_{o2} \tag{2.168}$$

where g_{mi} and r_{oi} are, respectively, the small-signal transconductance and output impedance for transistor M_i. An addition gain stage A_{dd}, as in Figure 2.93b (implemented in the regulated current mirror by M_{add}) increases the output impedance almost by a factor of $(A_{dd} + 1)$ and gives

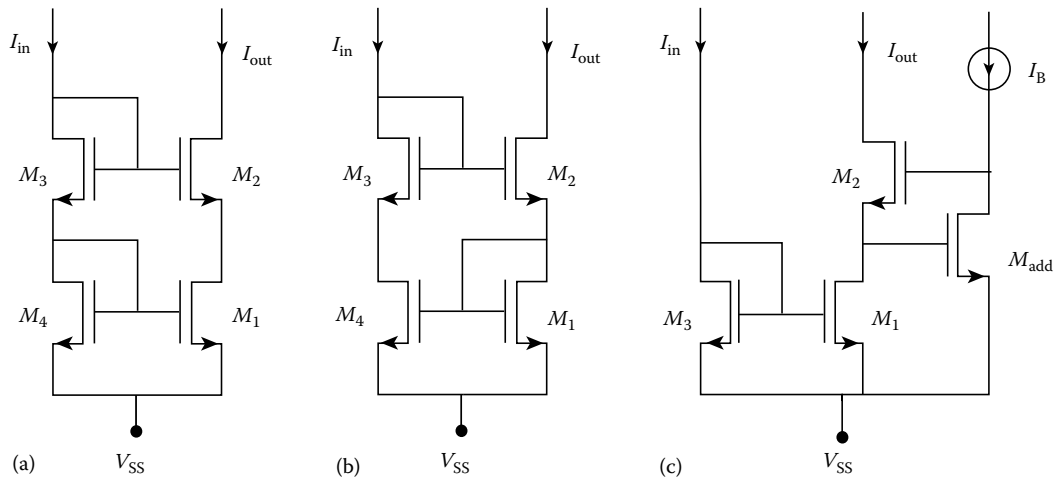


FIGURE 2.92 (a) Cascode, (b) improved Wilson, and (c) regulated current mirrors.

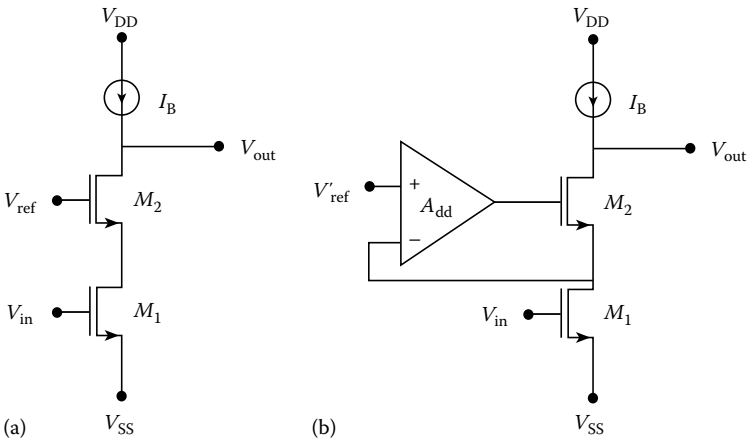


FIGURE 2.93 (a) Cascode stage and (b) cascode stage with an additional gain stage.

$$r_{o,b} = [g_{m2}r_{o2}(A_{dd} + 1) + 1]r_{o1} + r_{o2} \tag{2.169}$$

As a result, composite transistors with high output impedances can be obtained as shown in Figure 2.94 [10]. Figure 2.94a is directly obtained from the regulated current mirror, where M_{N1} and M_{N2} are cascoded and M_{P2} and M_{N4} compose an additional gain stage. The drain-source voltage of M_{N1} biased by I_1 is given by

$$V'_{DS,a} = \sqrt{\frac{2I_1}{K_4}} + V_T \tag{2.170}$$

Figure 2.94b, a modified version of (a), employs the biasing technique in Ref. [12], biasing V_{DS} for a triode-mode V–I converter. The resultant $V'_{Ds,b}$ is given by

$$V'_{DS,b} = \sqrt{\frac{2I_1}{K_4}} - \sqrt{\frac{2I_2}{K_5}} \tag{2.171}$$

Therefore, unlike the circuit shown in Figure 2.93a, where $V'_{Ds,a}$ is larger than V_T, M_{N1} can be biased at the edge of saturation by properly choosing currents I_1 and I_2 or K_4 and K_5 In addition, V_F provides a low impedance node for folded cascode configurations. Figure 2.94c, also called the super-MOS transistor [11], uses a similar concept. The CMOS cascode gain stage, composed by M_{P2}, M_{P4}, M_{N4}, and M_{N6}, gives a higher gain than the previous two circuits. Since M_{N8} in the series composite transistor (constituted by M_{N7} and M_{N8}) is always operating in the triode region, V_{DS} of M_{N8} can be very small, and M_{N1} can also be biased at the edge of saturation. However, due to its circuit complexity, the input range for V_{GS} is limited. The drain currents of the circuits shown in Figure 2.94b versus V_{DS} with various V_{GS} are compared to those obtained from a single transistor and are given in Figure 2.95. It can be seen that the output impedance of the composite transistor is significantly larger than that of a single one. The use of super MOS transistors in the design of high-gain operational amplifiers is discussed in Ref. [11].

2.4.5 Basic Voltage Gain Cells

In this subsection, we discuss simple voltage amplifier circuits implemented in NMOS, CMOS, and BiCMOS technologies.

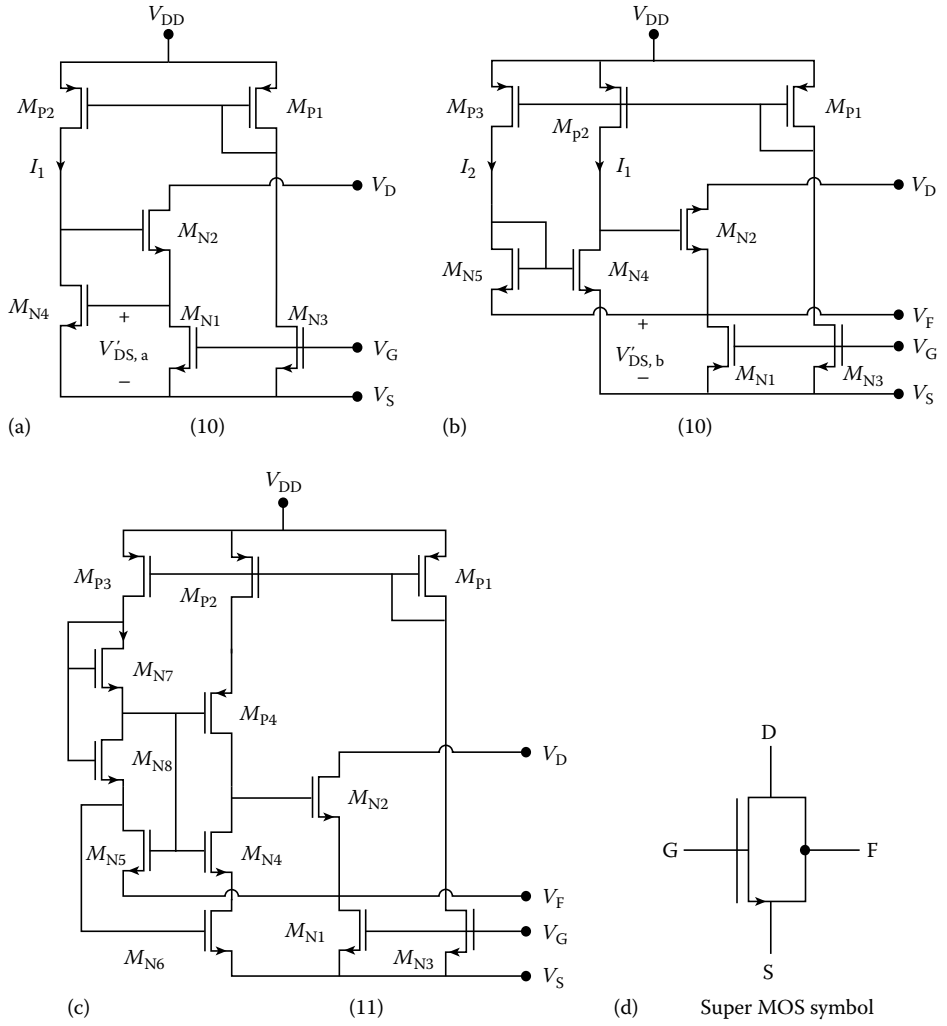


FIGURE 2.94 Composite super NMOS transistors. (a, b) (From Huang, S.-C., Systematic design solutions for analog VLSI circuits, PhD dissertation, Department of Electrical Engineering, Ohio State University, 1994.) (c) (From Bult, K. and Geelen, G.J., *J. Analog Integrat. Circuits Signal Process.*, 1, 119, 1991.)

2.4.5.1 NMOS Amplifier

Figure 2.96a shows an enhancement common-source NMOS amplifier with an enhancement load. M_1 is the driving (amplifying) transistor and the diode-connected transistor M_2 is the load device. The large-signal transfer characteristics of the amplifier is shown in Figure 2.96b and displays three well-defined regions. In region I, M_1 is off since $v_I < V_{T1}$. M_2, however, is always in the saturation region and is conducting a small current. The voltage across it is V_{T2} and hence the output voltage, v_o, is $V_{DD} - V_{T2}$. In region II, M_1 is conducting and is operating in saturation and the transfer curve is linear. Finally in region III, M_1 leaves the saturation region and enters the triode region. For the circuit to operate as an amplifier, the dc operating point must be located in the linear region (region II). Assuming that both M_1 and M_2 have the same threshold voltage V_T, but different values of K (K_1 and K_2) and neglecting both channel length modulation and body effects, we write

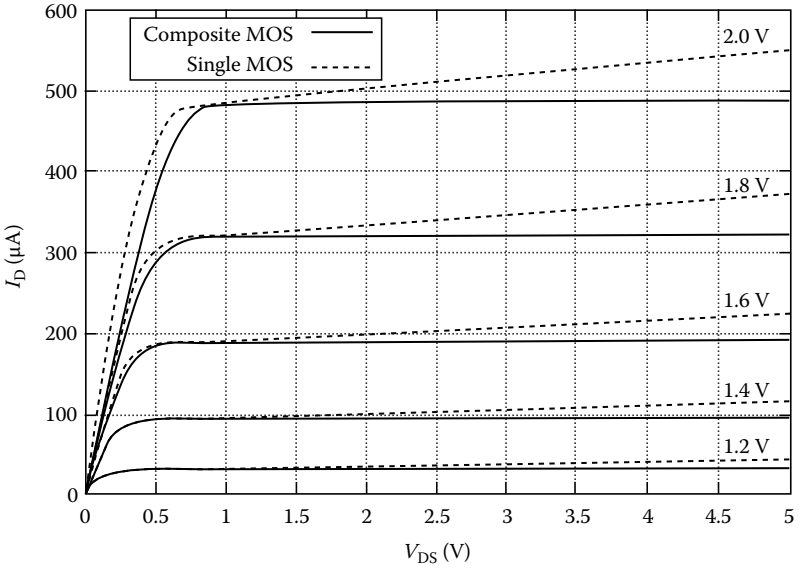


FIGURE 2.95 Simulated I_D curves for the high-output impedance composite transistor shown in Figure 2.94b and for a single transistor.

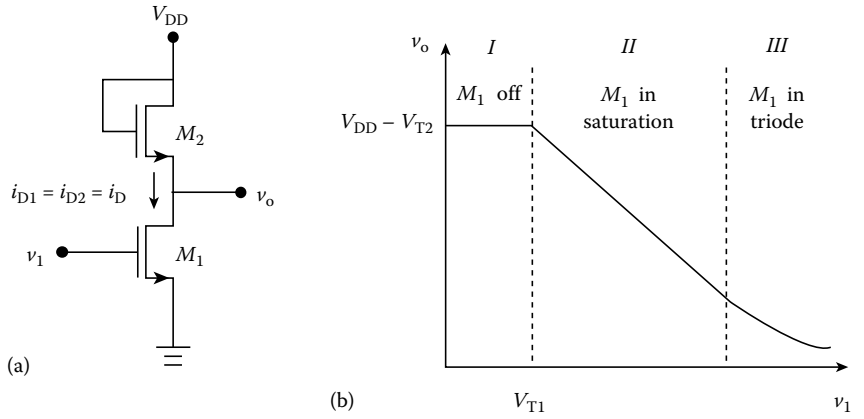


FIGURE 2.96 (a) The NMOS amplifier. (b) Transfer characteristics.

$$i_{D1} = i_{D2} = i_D = \frac{K_1}{2}(v_1 - V_T)^2 \tag{2.172}$$

and

$$\begin{aligned} i_D &= \frac{K_2}{2}(v_{GS2} - V_T)^2 \\ &= \frac{K_2}{2}(V_{DD} - v_o - V_T)^2 \end{aligned} \tag{2.173}$$

Combining Equations 2.172 and 2.173 and with some manipulations, we obtain

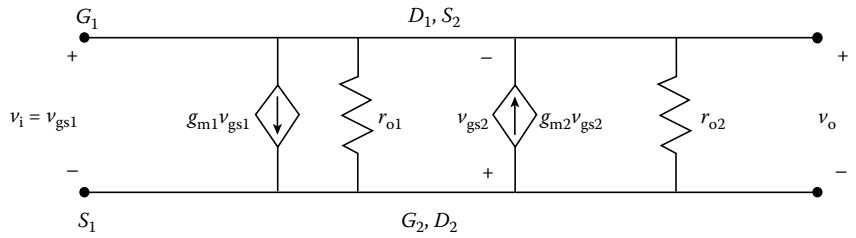


FIGURE 2.97 Small-signal equivalent circuit of the NMOS amplifier.

$$v_o = \left(V_{DD} - V_T + \sqrt{\frac{K_1}{K_2}}V_T\right) - \sqrt{\frac{K_1}{K_2}}v_I \tag{2.174}$$

which is a linear equation between v_o and v_I. This is obviously the equation of the straight-line portion (region II) of the transfer curve.

The first term in Equation 2.174 represents the dc component of the output voltage V_o. The second term represents the small-signal component and thus the ac small-signal gain of the amplifier A_v is

$$A_v = \frac{v_o}{v_i} = -\sqrt{\frac{K_1}{K_2}} = -\sqrt{\frac{(W/L)_1}{(W/L)_2}} \tag{2.175}$$

The small-signal equivalent circuit of the amplifier in Figure 2.96a is shown in Figure 2.97. Since D_2 and G_2 are connected in M_2, the voltage across the controlled current–source $g_{m2}v_{gs2}$ is v_{gs2}. Therefore, the controlled current–source can be represented by a resistance $1/g_{m2}$. Since $v_{gs1} = v_i$, we obtain the voltage gain as follows:

$$A_v = \frac{v_o}{v_i} = -\frac{g_{m1}}{g_{m2} + (1/r_{o1}) + (1/r_{o2})} \tag{2.176}$$

Now, if r_{o1} and r_{o2} are much larger than $(1/g_{m2})$, the gain reduces to $A_v \simeq -g_{m1}/g_{m2}$, which can easily be shown to lead to the expression in Equation 2.175. Note that the gain can also be determined by inspection from the circuit in Figure 2.96a as $-g_{m1}$ multiplied by the equivalent small-signal resistance seen at the drain of M_1, which is $(1/g_{m2}||r_{o1}||r_{o2})$.

Practically, M_1 and M_2 share the same substrate, which is normally connected to the most negative supply voltage in the circuit (ground in this case). It follows that M_2 suffers from body effect, which is modeled in the small-signal equivalent circuit by a controlled current-source $g_{mb2}v_{bs2}$ connected between the two output terminals in Figure 2.97, where $v_{bs2} = v_{gs2}$ and $g_{mb2} = \chi g_{m2}$ and χ is a function of the dc source–body voltage V_{SB} and lies in the range 0.1–0.3 [13]. Taking the body effect of M_2 into account, the amplifier gain becomes

$$A_v = -\frac{g_{m1}}{g_{m2} + g_{mb2}} = -\frac{g_{m1}}{g_{m2}}\frac{1}{1+\chi} \tag{2.177}$$

2.4.5.2 CMOS Amplifier

In CMOS technology, both n-channel and p-channel devices are available, and are usually fabricated in a way that eliminates the body effect. The basic CMOS amplifier is shown in Figure 2.98. Here, M_2 and M_3 in Figure 2.98c are a pair of PMOS devices operating as a current source active load and implement the

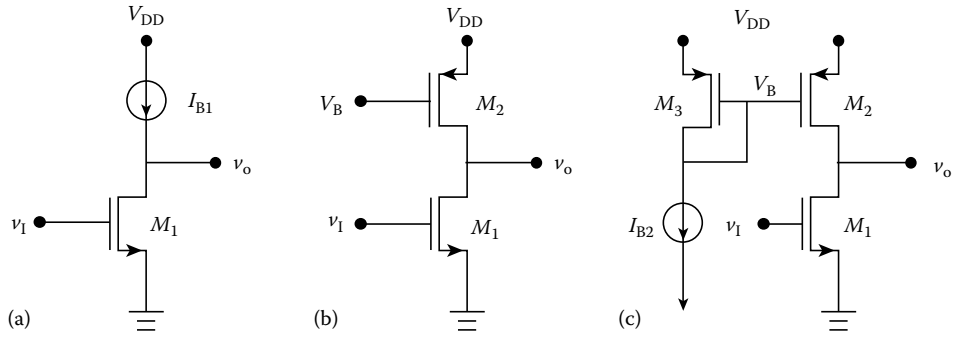


FIGURE 2.98 CMOS amplifier: (a) basic circuit, (b and c) CMOS implementations.

current source I_{B1} in Figure 2.98a. M_2 is biased in the saturation region and when M_1 is operating in the saturation region, the small-signal voltage gain will be equal to $-g_{m1}$, multiplied by the total resistance seen between the output and ground which is $(r_{o1}||r_{o2})$.

Cascode versions of the amplifier as shown previously in Figure 2.93 can be used to boost the gain significantly. For instance, the cascode amplifier in Figure 2.93a has a gain equal to the effective transconductance $-g_{meff}$, multiplied by $r_{o,a}$ given by Equation 2.168, where g_{meff} is given by [14]

$$g_{meff} = g_{m1}\frac{g_{m2}r_{o1} + (r_{o1}/r_{o2})}{g_{m2}r_{o1} + (r_{o1}/r_{o2}) + 1} \tag{2.178}$$

2.4.5.3 BiCMOS Amplifiers

A BiCMOS technology combines bipolar and CMOS transistors on a single substrate. A bipolar transistor has the advantage over an MOS of a much higher transconductance (g_m) for the same dc bias current. Also, bipolar transistors have better high-frequency performance than their MOS counterparts. On the other hand, the practically infinite input resistance at the gate of a MOSFET makes it possible to design amplifiers with extremely high input resistance and an almost zero input bias current. For these reasons, there has been an increasing interest in BiCMOS technologies for implementing high-performance integrated circuits. While most BiCMOS processes offer high-quality NMOS, PMOS, and NPN transistors, advanced BiCMOS processes offer PNP transistors as well.

Figure 2.99 shows three basic folded-cascode single-ended high-performance BiCMOS amplifiers [15]. The main features of these amplifiers are a high gain-bandwidth product and extremely high dc input

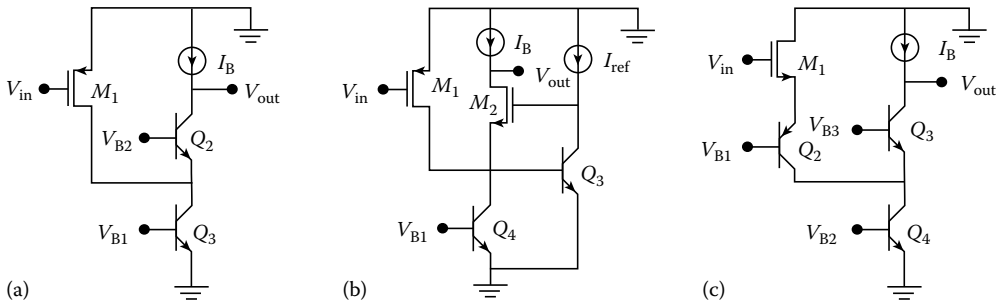


FIGURE 2.99 BiCMOS basic amplifier circuits: (a) common-source, common-base, (b) common-source, common-gate with active-feedback, and (c) common-drain, common-base, common-base.

impedance. The high gain is achieved by cascoding transistors in the signal path (M_1, Q_2 in Figure 2.99a, M_1, M_2 in Figure 2.99b, and M_1, Q_2, Q_3, in Figure 2.99c). The high bandwidth is achieved by exploiting the exponential nature of the current–voltage characteristics of bipolar transistors. For instance, let us consider the amplifier circuit in Figure 2.99a. The internal node in the signal path at the emitter of Q_2 has an extremely low impedance. This is due to the fact that while the emitter current can change significantly with the input signal, the emitter voltage remains almost constant. The same argument can be made about the internal nodes in the signal paths of the other two amplifiers (the base of Q_3 in Figure 2.99b and the emitters of Q_2 and Q_3 in Figure 2.99c). The low impedance of internal nodes in the signal path places nondominant poles at very high frequencies.

2.4.5.4 Differential Amplifier

The amplifier circuits discussed previously are of the single-ended type. A single-ended amplifier has both input and output voltage signals referred to ground. In most IC applications, a differential amplifier is utilized. In this case, the amplifier has a differential input and may also have a differential output, in which case it is called a fully differential amplifier. It is usually easy and straightforward to convert single-ended amplifiers to differential architectures.

The most widely used differential amplifier is based on the common-source or common-gate differential pairs shown respectively in Figure 2.81b and c. The common-source pair is shown here again in Figure 2.100. The only difference here is that the circuit is biased by a constant current source I that is usually implemented using a current-mirror circuit (see Figure 2.81d).

Assuming that M_1 and M_2 are identical and neglecting both channel length modulation and body effect, we write

$$i_{D1} = \frac{K}{2}(v_{GS1} - V_T)^2 \tag{2.179}$$

$$i_{D2} = \frac{K}{2}(v_{GS2} - V_T)^2 \tag{2.180}$$

Taking the square root of both sides in each of the two equations above and defining the differential input as $v_{id} = v_{GS1} - v_{GS2}$, we get

$$\sqrt{i_{D1}} - \sqrt{i_{D2}} = \sqrt{\frac{K}{2}}\, v_{id} \tag{2.181}$$

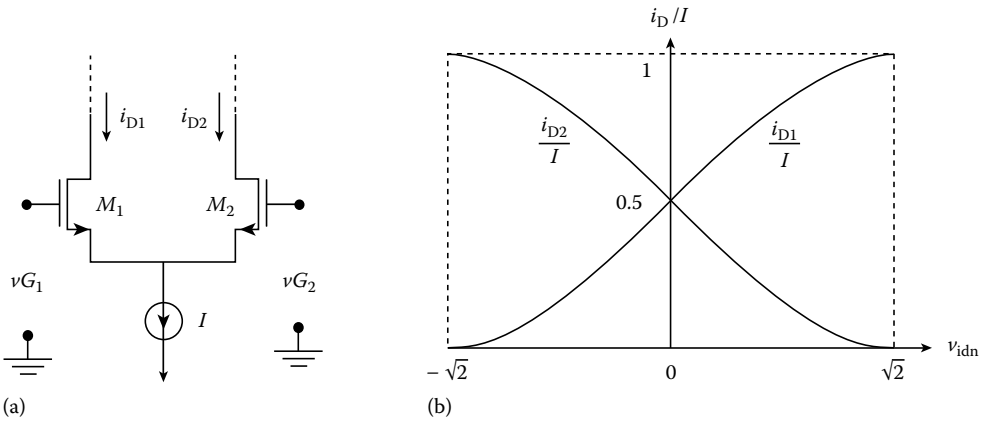


FIGURE 2.100 MOS differential pair: (a) the circuit and (b) the transfer characteristic.

But since the current–source bias imposes the constraint that $i_{D2} + i_{D2} = I$, one can easily show that

$$i_{D1,2} = \frac{I}{2} \pm \sqrt{KI}\left(\frac{v_{id}}{2}\right)\sqrt{1 - \frac{(v_{id}/2)^2}{(I/K)}} \tag{2.182}$$

At the bias point, the small-signal differential input voltage v_{id} is zero,

$$v_{GS1} = v_{GS2} = V_{GS} \quad \text{and} \quad i_{D1} = i_{D2} = I/2$$

This can be used to rewrite the preceding equation as follows:

$$i_{D1,2} = \frac{I}{2} \pm \left(\frac{I}{V_{GS} - V_T}\right)\left(\frac{v_{id}}{2}\right)\sqrt{1 - \left(\frac{v_{id}/2}{V_{GS} - V_T}\right)^2} \tag{2.183}$$

And for $v_{id}/2 \ll V_{GS} - V_T$ (small-signal approximation),

$$i_{D1,2} \simeq \frac{I}{2} \pm \left(\frac{I}{V_{GS} - V_T}\right)\left(\frac{V_{id}}{2}\right) \tag{2.184}$$

The differential pair tranconductance g_m, defined as $g_m = (i_{D1} - i_{D2})/v_{id}$ is then given by $I/(V_{GS} - V_T)$. We recall that a single MOS transistor biased at a drain current I_D has a transconductance $2I_D/(V_{GS} - V_T)$. Thus, we see that each transistor in the pair has a transconductance $2(I/2)/(V_{GS} - V_T)$, which is equal to the differential pair transconductance, g_m. Equations 2.183 and 2.184 indicate that for small-signal inputs, the current in M_1 increases by i_d and that in M_2 decreases by i_d. From Equation 2.183, we can find v_{id} at which current steering between M_1 and M_2 occurs that is $i_{D1} = I$ and $i_{D2} = 0$ or vice versa for negative v_{id}. Equating the second term in Equation 2.183 to $I/2$, we get

$$|v_{id}|\,\max = \sqrt{2}(V_{GS} - V_T) \tag{2.185}$$

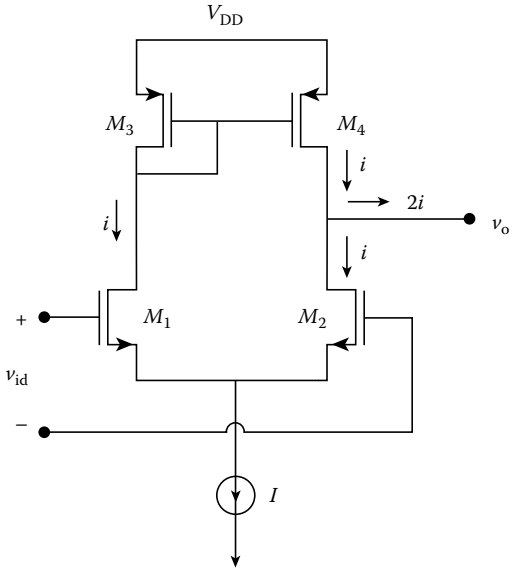


FIGURE 2.101 Simple differential-input, single-ended output CMOS amplifier.

Figure 2.100b shows plots of the normalized currents i_{D1}/I and i_{D2}/I versus the normalized differential input voltage $v_{idn} = v_{id}/(V_{GS} - V_T)$.

A simple CMOS differential amplifier is shown in Figure 2.101, where the PMOS pair is used as an active load. The small-signal current i is given by $g_m(v_{id}/2)$, where $g_m = I/(V_{GS} - V_T)$. The small-signal output voltage is given by $v_o = 2i(r_{o2}||r_{o4})$ and the voltage gain is $A_v = v_o/v_{id} = g_m((r_{o2}||r_{o4})$.

When $v_{id} = 0$, the bias current I does not actually split equally between M_1 and M_2. This is due to mismatches in K, ΔK, and V_T, ΔV_T, which contribute to a dc offset voltage that is usually larger than that in differential amplifiers implemented with bipolar transistors. For instance, modern silicon-gate MOS technologies have ΔV_T as high as 2 mV [13]. Note that ΔV_T has no counterpart in BJTs.

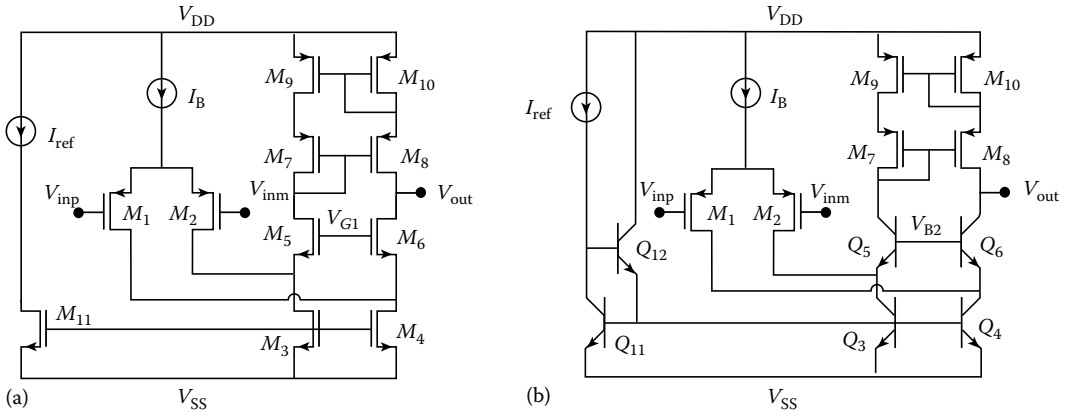


FIGURE 2.102 Folder-cascode amplifier: (a) common-source, common-gate and (b) common-source, common-base.

2.4.5.5 Folded-Cascode Operational Amplifier

The folded-cascode operational amplifier (op-amp) is a basic building block in modern analog-integrated circuits. Figure 2.102 shows two folded-cascode op-amp circuits in CMOS and BiCMOS technologies [15]. Actually, several combinations of bipolar and CMOS devices could be used in the design of this amplifier. Here, we assume that PNP bipolar transistors are not available, which implies a BiCMOS process having less complexity. The input common-source MOS pair is of the PMOS type. The cascode common-gate or common-base pair (M_5 and M_6 in Figure 2.102a and Q_5 and Q_6 in Figure 2.102b) is ''folded'' and, therefore, is implemented with devices of the opposite type to that used in the input pair. This is unlike the basic ''unfolded'' cascode amplifier in Figure 2.93, where both input and cascode devices are of the same type.

The greater values of transconductance associated with the common-base bipolar devices in the BiCMOS op-amp place the nondominant parasitic poles at much higher frequencies. Note that the BiCMOS op-amp circuit is based on the single-ended amplifier shown in Figure 2.99a. The BiCMOS op-amp combines the increased bandwidth the advantages of an MOS input stage; namely a nearly infinite input impedance, a zero input bias current, and a higher slew rate [13].

2.4.6 Conclusion

Analog design is more complicated and less systematic than digital design and involves many trade-offs to meet certain design specifications. It strongly relies on human heuristics. The transfer of these human experiences into a computer-aided design environment is essential to the success of analog design in the context of VLSI of both analog and mixed analog/digital integrated circuits. This transfer, however, requires the development of systematic approaches to the analysis and design of analog integrated circuits. To this end, understanding the basic operations of the analog cells discussed here is critical. The use of these cells in the systematic design of analog VLSI systems, such as filters and data converters, is discussed in Ref. [10].

References

1. M. Ismail and T. Fiez, *Analog VLSI: Signal and Information Processing*, New York: McGraw-Hill, 1994.
2. M. Ismail, S.-C. Huang, and S. Sakurai, Continuous-time signal processing, in *Analog VLSI: Signal and Information Processing*, M. Ismail and T. Fiez, Eds., New York: McGraw-Hill, 1994, Chapter 3.

3. Z. Wang, Analytical determination of output resistance and DC matching errors in MOS current mirrors, *IEE Proc.: Pt. G*, 137, 397–404, Oct. 1990.
4. M. C. H. Cheng and C. Toumazou, Linear composite MOSFETs (COMFETs), *Electron. Lett.*, 27, 1802–1802, Sept. 1991.
5. Y. Tsividis, M. Banu, and J. M. Khoury, Continuous-time MOSFET-C filters in VLSI, *IEEE J. Solid-State Circuits*, SC-21, 15–30, Feb. 1986.
6. M. Banu and Y. Tsividis, Floating voltage-controlled resistors in CMOS technology, *Electron. Lett.*, 18, 678, 679, July 1982.
7. M. Ismail, Four-transistor continuous-time MOS transconductor, *Electron. Lett.*, 23, 1099, 1100, Sept. 1987.
8. B. Nauta, A CMOS transconductance-C filter technique for very high frequencies, *IEEE J. Solid-State Circuits*, 27, 142–153, Feb. 1992.
9. J. Ramirez-Angulo, Applications of composite BiCMOS transistors, *Electron. Lett.*, 27, 2236–2238, Nov. 1991.
10. S.-C. Huang, Systematic design solutions for analog VLSI circuits, PhD dissertation, Department of Electrical Engineering, Ohio State University, Columbus, OH, 1994.
11. K. Bult and G. J. Geelen, The CMOS gain-boosting technique, *J. Analog Integrat. Circuits Signal Process.*, 1, 119–135, 1991.
12. U. Gatti, F. Maloberti, and G. Torelli, A novel CMOS linear transconductance cell for continuous-time filters, in *Proc. IEEE Int. Symp. Circuits Syst.* (ISCAS), 1990, pp. 1173–1176.
13. A. S. Sedra and K. C. Smith, *Microelectronic Circuits* (Series in Electrical Engineering), 3rd ed., Philadelphia: Holt, Rinehart & Winston, 1991, Chapters 5, 6, and 10.
14. K. Bult, Basic CMOS circuit techniques, in *Analog VLSI: Signal and Information Processing*, M. Ismail and T. Fiez, Eds., New York: McGraw-Hill, 1994, Chapter 2.
15. S. R. Zarabadi, M. Ismail, and F. Larsen, Basic BiCMOS circuit techniques, in *Analog VLSI: Signal and Information Processing*, M. Ismail and T. Fiez, Eds., New York: McGraw-Hill, 1994, Chapter 5.

3

High-Performance Analog Circuits

3.1 Broadband Bipolar Networks **3**-1
Introduction • Miller's Theorem • Bipolar Transistor Modeling at High Frequencies • Single-Gain Stages • Neutralization of C_μ • Negative Feedback • RF Bipolar Transistor Layout • Bipolar Current-Mode Broadband Circuits • Broadband Amplifier Stability • Conclusions
Appendix A: Transfer Function and Bandwidth Characteristic of Current-Feedback **3**-41
Appendix B: Transfer Function and Bandwidth Characteristic of Voltage-Feedback **3**-43
Appendix C: Transconductance of the Current-Feedback Op-Amp Input Stage **3**-44
Appendix D: Transfer Function of Widlar Current Mirror **3**-45
Appendix E: Transfer Function of Widlar Current Mirror with Emitter Degeneration Resistors **3**-47
References **3**-47
3.2 Bipolar Noise **3**-48
Thermal Noise • Shot Noise • Generation–Recombination Noise • *1/f* Noise • Noise $1/f^2$ • Burst Noise—RTS Noise • Avalanche Noise • Noise Characterization
References **3**-55

Chris Toumazou
Imperial College of Science, Technology, and Medicine

Alison Payne
Imperial College of Science, Technology, and Medicine

John Lidgey
Oxford Brookes University

Alicja Konczakowska
Gdańsk University of Technology

Bogdan M. Wilamowski
Auburn University

3.1 Broadband Bipolar Networks

Chris Toumazou, Alison Payne, and John Lidgey

3.1.1 Introduction

Numerous textbooks have presented excellent treatments of the design and analysis of broadband bipolar amplifiers. This chapter is concerned with techniques for integrated circuit amplifiers, and is written mainly as a tutorial aimed at the practicing engineer.

For broadband polar design, it is first important to identify the key difference between lumped and distributed design techniques. Basically when the signal wavelengths are close to the dimensions of the integrated circuit, then characteristic impedances become significant, lines become lossy, and we essentially need to consider the circuit in terms of transmission lines. At lower frequencies where the signal

wavelength is much larger than the dimensions of the circuit, the design can be considered in terms of lumped components, allowing some of the more classical low-frequency analog circuit techniques to be applied. At intermediate frequencies, we enter the realms of hybrid lumped/distributed design. Many radio-frequency (RF) designs fall into this category, although every day we see new technologies and circuit techniques developed that increase the frequency range for which lumped approaches are possible. In broadband applications, integrated circuits (ICs) are generally designed without the use of special microwave components, so broadband techniques are very similar to those employed at lower frequencies. However, several factors still have to be considered in RF design: all circuit parasitics must be identified and included to ensure accurate simulation; feedback can generally only be applied locally as phase shifts per stage are significant; the cascading of several local feedback stages is difficult since alternating current (ac) coupling is often impractical; the NPN bipolar transistor is the main device used in silicon, since it has potentially a higher f_t than PNP bipolar or MOSFET devices; active PNP loads are generally avoided due to their poor frequency and noise performance and so resistive loads are used instead.

The frequency performance of an RF or broadband circuit will depend on the frequency capability of the devices used, and no amount of good design can compensate for transistors with an inadequate range. As a rule, designs are kept as simple as possible, since at high frequencies all components have associated parasitics.

3.1.2 Miller's Theorem

It is important to describe at the outset a very useful approximation that will assist in simplifying the high-frequency analysis of some of the amplifiers to be described. The technique is known as Miller's theorem and will be briefly discussed here. A capacitor linking input to output in an inverting amplifier results in an input-referred shunt capacitance that is multiplied by the voltage gain of the stage, as shown in Figure 3.1. This increased input capacitance is known as the Miller capacitance.

It is straightforward to show that the input admittance looking into the inverting input of the amplifier is approximately $Y_{in} = j\omega C_f(1 + A)$. The derivation assumes that the inherent poles within the amplifier are at a sufficiently high frequency so that the frequency response of the circuit is dominated by the input of the amplifier. If this is not the case, then Miller's approximation should be used with caution as it will be discussed later. From the preceding model, it is apparent that the Thévenin input signal sources see an enlarged capacitance to ground. Miller's approximation is often a useful way of simplifying circuit analysis by assuming that the input dominant frequency is given by the simple low-pass *RC* filter in Figure 3.1. However, the effect is probably one of the most detrimental in broadband amplifier design, affecting both frequency performance and/or stability.

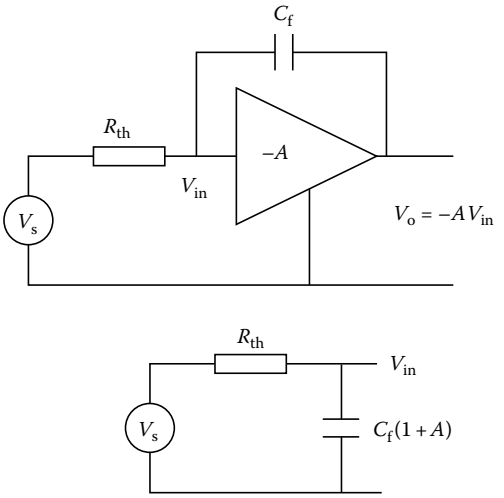


FIGURE 3.1 Example of the Miller effect.

3.1.3 Bipolar Transistor Modeling at High Frequencies

In this section, we consider the high-frequency small-signal performance of the bipolar transistor. The section assumes that the reader has some knowledge of typical device parameters, and has some familiarity with the technology. For small-signal analysis, the simplified hybrid-π model shown in Figure 3.2 is used,

where

r_b is the base series resistance
r_c is the collector series resistance
r_π is the dynamic base–emitter resistance
r_o is the dynamic collector–emitter resistance

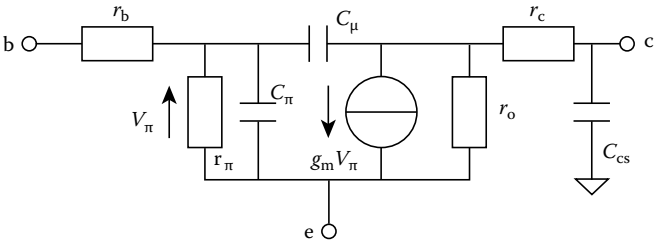


FIGURE 3.2 Hybrid π model of BJT.

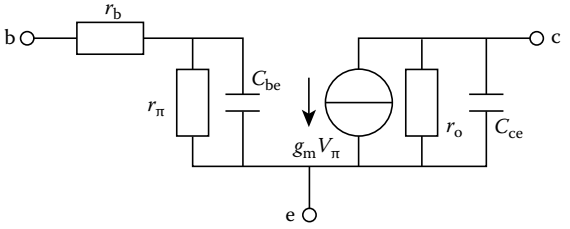


FIGURE 3.3 Simplified Miller-approximated hybrid π model of BJT.

C_π is the base–emitter junction capacitance
C_μ is the collector–base junction capacitance
C_{cs} is the collector–substrate capacitance
g_m is the small-signal transconductance

At low frequencies, the Miller approximation allows the hybrid-π model to be simplified to the circuit shown in Figure 3.3, where the net input capacitance now becomes $C_{be} = C_\pi + C_\mu (1 - A_v)$, the net output capacitance becomes $C_{ce} = C_\mu(1 - 1/A_v)$, where A_v is the voltage gain given by $A_v = (V_{ce}/V_{be}) \approx -g_m R_1$ where R_1 is the collector load resistance. r_c and C_{cs} have been neglected. Thus, $C_{be} \approx C_\pi + g_m R_1 C_\mu$ and $C_{ce} \approx C_\mu$. The output capacitance C_{ce} is often neglected from the small-signal model. The approximation $A_v = -g_m R_1$ assumes that $r_\pi \gg r_b$, and that the load is purely resistive. At high frequencies, however, we cannot neglect the gain roll-off due to C_π and C_μ, and even at frequencies as low as 5% of f_t the Miller approximation can introduce significant errors.

A simplified hybrid-π model that takes the high-frequency gain roll-off into account is shown in Figure 3.4. C_μ is now replaced by an equivalent current source $sC_\mu(V_\pi - V_{ce})$.

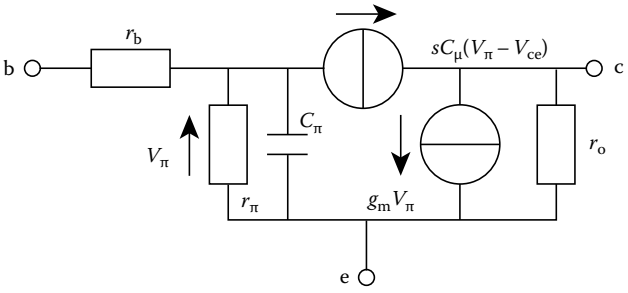


FIGURE 3.4 Simplified high-frequency model.

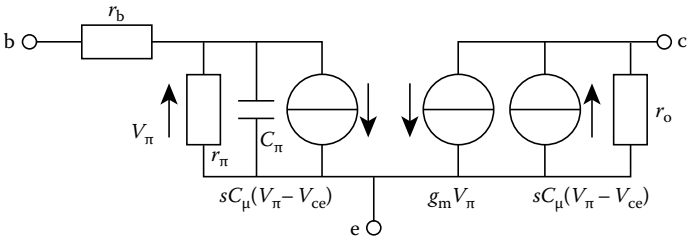


FIGURE 3.5 Split current sources.

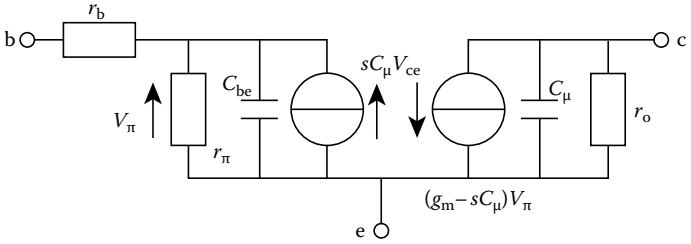


FIGURE 3.6 Modified equivalent circuit.

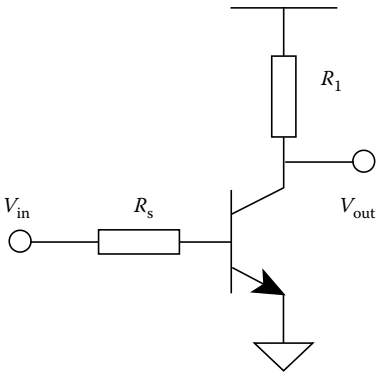


FIGURE 3.7 CE amplifier.

A further modification is to split the current source between the input and output circuits as shown in Figure 3.5.

Finally, the input and output component terms can be rearranged leading to the modified equivalent circuit shown in Figure 3.6, which is now suitable for broadband design. From Figure 3.6, the transconductance $(g_m - sC_\mu)$ shows the direct transmission of the input signal through C_μ. The input circuit current source $(sC_\mu V_{ce})$ shows the feedback from the output to the input via C_μ. Depending on the phase shift between V_{ce} and V_{be}, this feedback can cause high-frequency oscillation. At lower frequencies, $sC_\mu \ll g_m$ and $V_{ce}/V_\pi \approx -g_m R_1$, which is identical to the Miller approximation. The model of Figure 3.6 is the most accurate for broadband amplifier design, particularly at high frequencies.

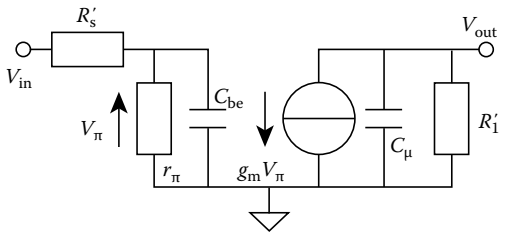


FIGURE 3.8 High-frequency model of the CE.

3.1.4 Single-Gain Stages

Consider now the high-frequency analysis of single-gain stages.

3.1.4.1 Common-Emitter (CE) Stage

Figure 3.7 shows a CE amplifier with load R_1 and source R_s. External biasing components are excluded from the circuit.

First analysis using the Miller approximation yields the small-signal high-frequency model shown in Figure 3.8,

where

$$\begin{aligned} &R_{1'} = (R_1 \| r_0), \quad R_{s'} = R_s + r_b \quad \text{and} \quad C_{be} = C_\pi + g_m R_{1'} C_\mu \\ &\frac{V_\pi}{V_{in}} = \left(\frac{r_\pi}{r_\pi + R_{s'}}\right)\left(\frac{1}{1 + s(r_\pi \| R_{s'})C_{be}}\right) \\ &\frac{V_{out}}{V_\pi} = \frac{-g_m R_{1'}}{1 + sC_\mu R_{1'}} \end{aligned} \tag{3.1}$$

and thus

$$\frac{V_{out}}{V_{in}} = -\left(\frac{g_m R_{1'} r_\pi}{r_\pi + R_{s'}}\right)\left(\frac{1}{(1 + sC_\mu R_{1'})(1 + s(r_\pi \| R_{s'})C_{be})}\right) \tag{3.2}$$

This approximate analysis shows

- "Ideal" voltage gain $= -g_m R_1$
- Input attenuation caused by $R_{s'}$ in series with r_π
- Input circuit pole p_1 at $s = 1/C_{be}(r_\pi/R_{s'}) \approx 1/C_{be}R_{s'}$
- Output attenuation caused by r_0 in parallel with R_1
- Output circuit pole p_2 at $s = 1/C_\mu R_{1'}$

The input circuit pole is generally dominant, and thus the output pole p_2 can often be neglected. With a large load capacitance C_1, $p_2 \approx 1/C_1 R_{1'}$, and the gain and phase margin will be reduced. However, under these conditions the Miller approximation will no longer be valid, since the gain roll-off due to the load capacitance is neglected.

If we now consider analysis using the broadband hybrid-π model of Figure 3.6, then the equivalent model of the CE now becomes that shown in Figure 3.9, where

$$C_{be} = C_\pi + C_\mu, \quad R_{s'} = R_s + r_b \quad \text{and} \quad R_{1'} = R_1 \| r_0$$

From the model, it can be shown that

$$\frac{V_{out}}{V_\pi} = \frac{-(g_m - sC_\mu)R_{1'}}{1 + sC_\mu R_{1'}} \tag{3.3}$$

$$(V_{in} - V_\pi)/R_{s'} + sC_\mu V_{out} = V_\pi/r_\pi + sC_{be}V_\pi \tag{3.4}$$

and

$$V_{in} r_\pi + V_{out} sC_\mu r_\pi R_{s'} = V_\pi(r_\pi + R_{s'})(1 + sC_{be}(r_\pi \| R_{s'})) \tag{3.5}$$

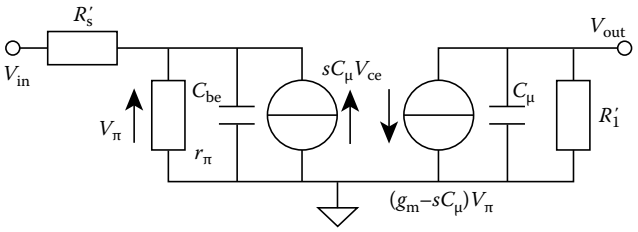


FIGURE 3.9 Equivalent circuit model of the CE.

Rearranging these equations yields

$$\frac{V_{out}}{V_{in}} = -\left(\frac{g_m R_{1'} r_\pi}{r_\pi + R_{s'}}\right) \times \left(\frac{1 - sC_\mu/g_m}{(1 + sC_{be}R_{s'})(1 + sC_\mu R_{1'}) + sC_\mu g_m R_{1'} R_{s'} - s^2 C_\mu^2 R_{s'} R_{1'}}\right) \tag{3.6}$$

This analysis shows that there is a right-hand-plane (RHP) zero at $s = 1/(C_\mu r_e)$, which is not predicted by the Miller approximation. Assuming $R_\pi \gg R_{s'}$ and $C_\pi \gg C_\mu$, the denominator can be written as

$$1 + s\left(R_{s'}\left(C_\pi + C_\mu g_m R_{1'}\right) + C_\mu R_{1'}\right) + s^2 C_\mu C_\pi R_{1'} R_{s'} \tag{3.7}$$

which can be described by the second-order characteristic equation

$$1 + s(1/p_1 + 1/p_2) + s^2/p_1 p_2 \tag{3.8}$$

By comparing coefficients in Equations 3.7 and 3.8, the sum of the poles is the same as that obtained in Equation 3.2 using the Miller approximation, but the pole product $p_1 p_2$ is greater. This means that the poles are farther apart than predicted by the Miller approximation. In general, the Miller approximation should be reserved for analysis at frequencies of operation well below f_t, and for situations where the capacitive loading is not significant. The equivalent circuit of Figure 3.9 therefore gives a more accurate result for high-frequency analysis. For a full understanding of RF behavior, computer simulation of the circuit including all parasitics is essential.

Since the CE stage provides high current and voltage gain, oscillation may well occur. Therefore, care must be taken during layout to minimized parasitic coupling between the input and output. The emitter should be at ground potential for ac signals, and any lead inductance from the emitter to ground will generate phase-shifted negative feedback to the base, which can result in instability.

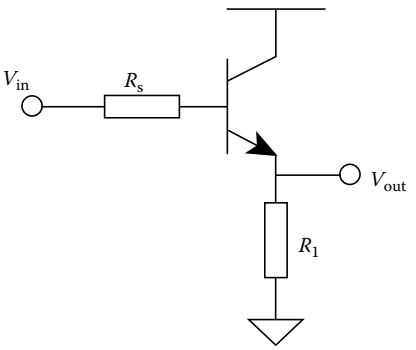


FIGURE 3.10 Common-collector amplifier.

3.1.4.2 Common-Collector (CC) Stage

The CC or emitter follower shown in Figure 3.10 is a useful circuit configuration since it generally serves to isolate a high-gain stage from a load. The high-frequency performance of this stage must be good enough not to degrade the frequency performance or stability of the complete amplifier. An equivalent high-frequency small-signal model of the CC is shown in Figure 3.11.

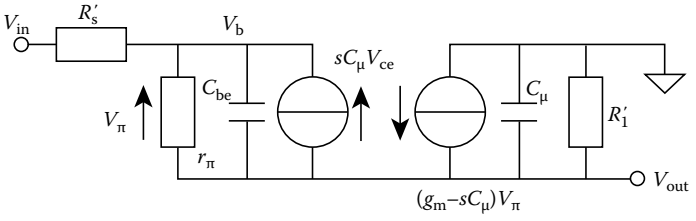


FIGURE 3.11 Equivalent circuit of the CC.

The following set of equations can be derived from Figure 3.11:

$$(V_{\text{in}} - V_{\text{b}})/R_{s'} = V_{\pi}/r_{\pi} + sC_{\text{be}}V_{\pi} + sC_{\mu}V_{\text{out}}, \quad V_{\text{b}} = V_{\text{out}} + V_{\pi} \tag{3.9}$$

and

$$V_{\pi}/r_{\pi} + sC_{\text{be}}V_{\pi} + sC_{\mu}V_{\text{out}} + \left(g_{\text{m}} - sC_{\mu}\right)V_{\pi} - sC_{\mu}V_{\text{out}} - V_{\text{out}}/R_{1'} = 0 \tag{3.10}$$

Rearranging these equations yields

$$\frac{V_{\text{out}}}{V_{\text{in}}} = \frac{R_{1'}(1 + g_{\text{m}}r_{\pi} + sC_{\pi}r_{\pi})}{(R_{s'} + r_{\pi})(1 + s(R_{s'} \| r_{\pi})C_{\text{be}}) + R_{1'}\left(1 + sC_{\mu}R_{s'}\right)(1 + g_{\text{m}}r_{\pi} + sC_{\pi}r_{\pi})} \tag{3.11}$$

The preceding expression can be simplified by assuming $R_{\pi} \gg R_{s'}$, $g_{\text{m}}r_{\pi} \gg 1$, $C_{\pi} \gg C_{\mu}$ to,

$$\frac{V_{\text{out}}}{V_{\text{in}}} = \left(\frac{r_{\pi}}{r_{\pi} + R_{s'}}\right)\left(\frac{1 + sC_{\pi}/g_{\text{m}}}{\left(1 + sC_{\mu}R_{s'}\right)(1 + sC_{\pi}/g_{\text{m}}) + (1 + sC_{\pi}R_{s'})/g_{\text{m}}R_{1'}}\right) \tag{3.12}$$

This final transfer function indicates the presence of a left-half-plane zero at $s = (g_{\text{m}}/C_{\pi}) = \omega_{\text{t}}$. The denominator can be rewritten as approximately

$$(1 + 1/g_{\text{m}}R_{1'}) + s\left(C_{\mu}R_{s'} + C_{\pi}/g_{\text{m}} + C_{\text{be}}R_{s'}/g_{\text{m}}R_{1'}\right) + s^2C_{\mu}C_{\pi}R_{s'}/g_{\text{m}} \tag{3.13}$$

which simplifies to

$$1 + s\left(C_{\pi}r_{\text{e}} + C_{\mu}R_{s'} + \left(C_{\mu} + C_{\pi}\right)r_{\text{e}}R_{s'}/R_{1'}\right) + s^2C_{\mu}C_{\pi}R_{s'}R_{1'} \tag{3.14}$$

Assuming a second-order characteristic form of $1 + s(1/p_1 + 1/p_2) + s^2/p_1p_2$, if $p_1 \ll p_2$, the above reduces to $1 + s/p_1 + s^2/p_1p_2$. If $(R_{s'}/R_{1'}) \ll 1$, then $p_1 \approx 1/(C_{\pi}r_{\text{e}})$, and this dominant pole will be approximately canceled by the zero. The frequency response will then be limited by the nondominant pole $p_2 \approx 1/C_{\mu}R_{s'}$.

The frequency response of a circuit containing several stages is thus rarely limited by the CC stage, due to this dominant pole-zero cancellation. For this analysis to be valid, $R_{s'} \ll R_{1'}$. As $R_{s'}$ increases the poles will move closer together, and the pole-zero cancellation will degrade. In practice, the CC stage is often used as a buffer, and is thus driven from a high source resistance into a low value load resistance.

A very important parameter of the common-collector stage is output impedance. It is generally assumed that the output impedance of a CC is low, also that there is good isolation between a load and the amplifying stage, and that any amount of current can be supplied to the load. Furthermore, it is assumed that capacitive loads will not degrade the frequency performance since the load will be driven by an almost short circuit. While this may be the case at low frequencies, it is a different story at high frequencies. Consider the following high-frequency analysis. We first assume that the small-signal model shown in Figure 3.12 is valid.

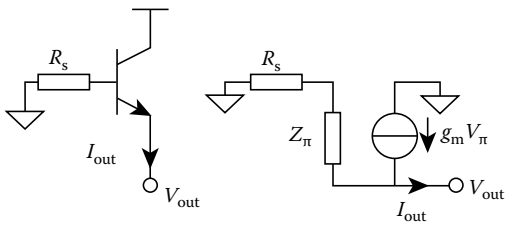


FIGURE 3.12 Equivalent circuit of the CC output stage.

From the Figure 3.12, the output impedance can be approximated as

$$\frac{V_{\text{out}}}{I_{\text{out}}} = \frac{Z_{\pi} + R_{s'}}{1 + g_{\text{m}}Z_{\pi}} \tag{3.15}$$

where $Z_\pi = (r_\pi \| C_{be})$ and $R_{s'} = R_s + r_b$. At very low frequencies ($\omega \to 0$):

$$R_{out} = \frac{r_\pi + R_{s'}}{1 + g_m r_\pi} \approx 1/g_m + R_{s'}/g_m r_\pi \approx r_e + R_{s'}/\beta \tag{3.16}$$

At very high frequencies ($\omega \to \infty$):

$$R_{out} = \frac{1/sC_{be} + R_{s'}}{1 + g_m/sC_{be}} \approx R_{s'} \tag{3.17}$$

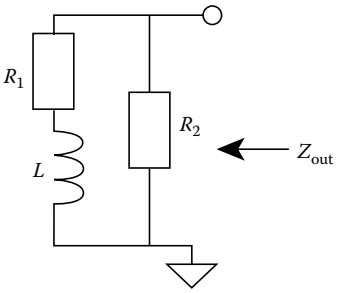


FIGURE 3.13 Equivalent high-frequency model of CC output stage.

If $r_e > R_{s'}$, then the output impedance decreases with frequency, that is, Z_{out} is capacitive. If $R_{s'} > r_e$, then Z_{out} increases with frequency, and so Z_{out} appears inductive. It is usual for an emitter follower to be driven from a high source resistance, thus the output impedance appears to be inductive and can be modeled as shown in Figure 3.13, where

$$R_1 = r_e + R_{s'}/\beta, \quad R_2 = R_{s'}, \quad L = R_{s'}/\omega_t$$

The inductive behavior of the CC stage output impedance must be considered in broadband design since any capacitive loading on this stage could result in peaking or instability. The transform from base resistance to emitter inductance arises because of the 90° phase shift between base and emitter currents at high frequencies, due principally to C_π. This transform property can be used to advantage to simulate an on-chip inductor by driving a CC stage from a high source resistance. Similarly, by loading the emitter with an inductor, we can increase the effective base series resistance $R_{s'}$ without degrading the noise performance of the circuit. A capacitive load will also be transformed by 90° between the base and emitter; for example, a capacitive loading on the base can look like a negative resistance at the emitter.

3.1.4.3 Common-Base (CB) Stage

The CB amplifier shown in Figure 3.14 offers the highest frequency performance of all the single-stage amplifiers. When connected as a unity gain current buffer, the CB stage operates up to the f_t of the transistor.

Using the simplified hybrid π model of Figure 3.3, it follows that

$$\frac{I_{out}}{I_{in}} \approx \frac{\beta}{\beta + 1} \quad \text{where} \quad \beta = \frac{\beta_o}{1 + s/\omega_o} \tag{3.18}$$

$$\frac{I_{out}}{I_{in}} \approx \frac{a_o}{1 + s/\omega_t} \quad \text{where} \quad a_o = \beta_o/(\beta_o + 1) \quad \text{and} \quad \omega_t = \beta_o \omega_o \tag{3.19}$$

The CB stage thus provides wideband unity current gain. Note that the input impedance of the CB stage is the same as the output impedance of the CC stage, and thus can appear inductive if the base series resistance is large.

In many situations, the CB stage is connected as a voltage amplifier, an example of this being the current-feedback amplifier, which will be discussed in a later section. Consider the following high-frequency analysis of the CB stage being employed as a voltage gain amplifier. Figure 3.15 shows the circuit together with a simplified small-signal model. From the equivalent model, the gain of the circuit can be approximated as

$$\frac{V_{out}}{V_{in}} = \frac{kR_1}{R_s}\left(\frac{1 - sC_\mu/g_m}{1 + s(C_\pi/g_m)(kR_{s'}/R_s)}\right) \tag{3.20}$$

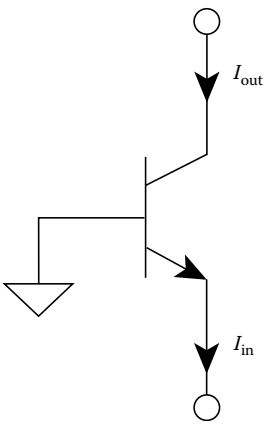


FIGURE 3.14 CB configuration.

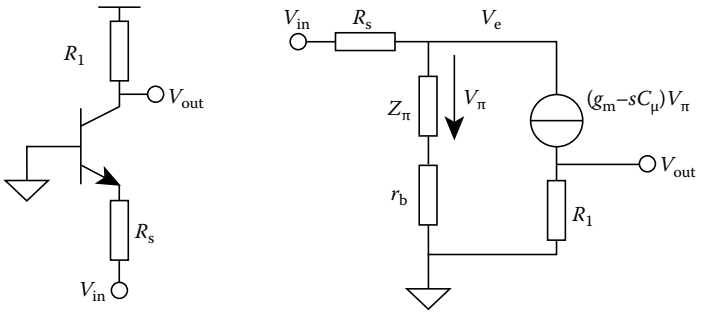


FIGURE 3.15 CB stage as a voltage amplifier.

where

$$R_{s'} = R_s + r_b, \quad \text{and} \quad k \approx \frac{R_s}{R_s + 1/g_m}$$

If $R_s \gg 1/g_m$, then $k \approx 1$ and so

$$\frac{V_{out}}{V_{in}} = \frac{R_1}{R_s}\left(\frac{1 - sC_\mu/g_m}{1 + s(C_\pi/g_m)(1 + r_b/R_s)}\right) \tag{3.21}$$

Thus, it can be seen that the circuit has an RHP zero at $s = 1/(r_e\ C_\mu)$, since $r_e = 1/g_m$ and a pole at $1/C_\pi r_e(1 + r_b/R_s) = \omega_t/(1 + r_b/R_s)$. Note that in the case of a current source drive ($R_s \gg r_b$), the pole is at the ω_t of the transistor. However, this does assume that the output is driven into a short circuit. Note also that there is an excellent isolation between the input and output circuits, since there is no direct path through C_μ and so no Miller effect.

3.1.5 Neutralization of C_μ

Many circuit techniques have been developed to compensate for the Miller effect in amplifiers and hence extend the frequency range of operation. The CE stage provides the highest potential power gain, but the bandwidth of this configuration is limited since the amplified output voltage effectively appears across the collector–base junction capacitance resulting in the Miller capacitance multiplication effect. This bandwidth limiting due to C_μ can be overcome by using a two-transistor amplifying stage such as the CE–CB cascode stage or the CC–CE cascade. Consider now a brief qualitative description of each in turn. The circuit diagram of the CE–CB cascode is shown in Figure 3.16.

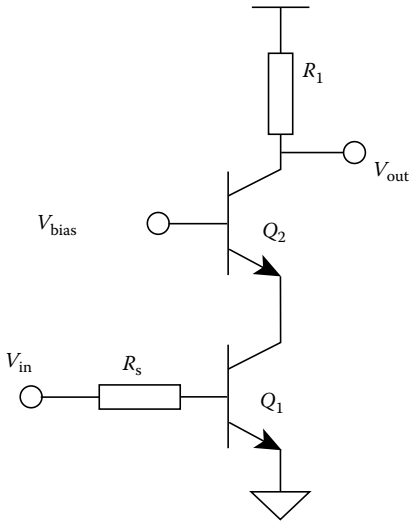


FIGURE 3.16 CE–CB cascode.

The CE transistor Q_1 provides high current gain of approximately β and a voltage gain of $A_{v1} \approx -g_{m1}R_1 = -g_{m1}r_{e2}$, which in magnitude will be close to unity. Therefore, the Miller multiplication of C_μ is minimized, and the bandwidth of Q_1 is maximized. The CB transistor Q_2 provides a voltage gain $A_{v2} \approx R_1/r_{e2}$. The total voltage gain of the circuit can be approximated as $A_v \approx -g_{m1}R_1$, which is equal to that of a single CE stage. The total frequency response is given by the cascaded response of both stages. Since both transistors exhibit wideband operation, then the dominant poles of each stage may be close in frequency. As a result, the total phase shift through the cascode

configuration is likely to be greater than that obtained with a single device, and care should be taken when applying negative feedback around the pair.

Consider now the CC–CE stage of Figure 3.17. In this case, voltage gain is provided by the CE stage transistor Q_2 and is $A_{v2} \approx -g_{m2}R_1$. This transistor is being driven from the low-output impedance of Q_1 and so the input pole frequency of this device ($\approx 1/C_{be2}R_{s2}$) is maximized. The CC stage transistor Q_1 is effectively a buffer that isolates C_μ of Q_2 from the source resistance R_s. The low-frequency voltage gain of this circuit is reduced when compared with a single-stage configuration because the input signal effectively appears across two base–emitter junctions.

The two-transistor configurations help to maintain a wideband frequency response by isolating the input and output circuits. In integrated circuit design, another method of neutralizing the effect of C_μ is possible when differential gain stages are used.

For example, Figure 3.18 shows a section of a differential input amplifier. If the inputs are driven differentially, then the collector voltages V_{c1} and V_{c2} will be 180° out of phase. The neutralization capacitors C_n thus inject a current into the base of each transistor that is equal and opposite to that caused by the intrinsic capacitance C_μ. Consequently, the neutralization capacitors should be equal to C_μ in order to provide good signal cancellation, and so they may be implemented from the junction capacitance of two dummy transistors with identical geometries to Q_1 and Q_2 as shown in Figure 3.19.

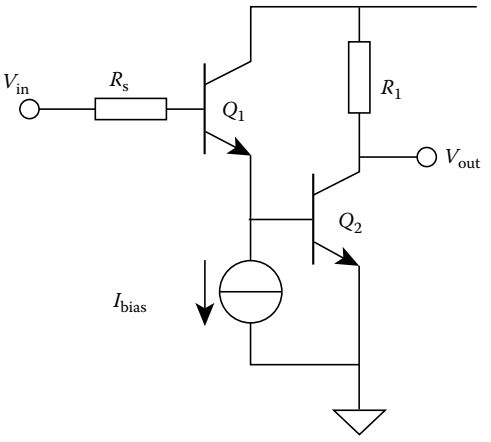


FIGURE 3.17 CC–CE stage.

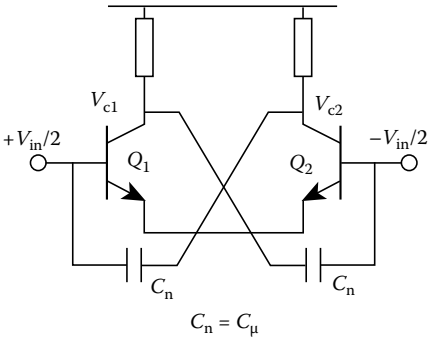


FIGURE 3.18 Differential gain stage.

3.1.6 Negative Feedback

Negative feedback is often employed around high-gain stages to improve the frequency response. In effect, the gain is reduced in exchange for a wider, flatter bandwidth. The transfer function of a closed-loop system can be written

$$H(s) = \frac{A(s)}{1 + A(s)B(s)} \tag{3.22}$$

where $A(s)$ is the open-loop gain and $B(s)$ is the feedback fraction. If the open-loop gain $A(s)$ is large, then $H(s) \approx 1/B(s)$. In RF design, compound or cascaded stages can produce excessive phase shifts that result in instability when negative feedback is applied. To overcome this problem, it is generally accepted to apply local negative feedback around a

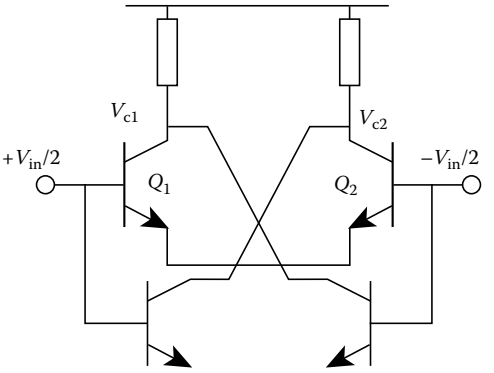


FIGURE 3.19 Implementation of neutralization capacitors.

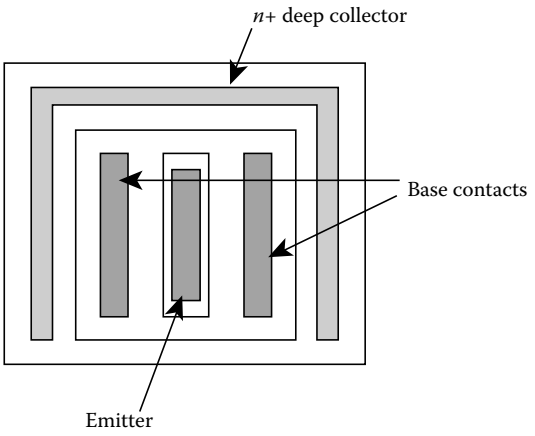


FIGURE 3.20 Stripe geometry.

single stage only. However, the open-loop gain of a single stage is usually too low for the approximation $H(s) = 1/B(s)$ to hold.

3.1.7 RF Bipolar Transistor Layout

When laying out RF transistors, the aim is to

- Minimize C_μ and C_π
- Minimize base width to reduce the forward transit time t_∂ and thus maximize ∂_t
- Minimize series resistance r_b and r_c

To minimize junction capacitance, the junction area must be reduced; however, this will tend to increase the series resistance. Transistors are generally operated at fairly high currents to maximize ∂_t. However, if the emitter gets too crowded, then the effective value of β will be reduced. The requirements given above are generally best met by using a stripe geometry of the type shown in Figure 3.20.

The stripe geometry maximizes the emitter area-to-periphery ratio, which reduces emitter crowding while minimizing the junction capacitance. The length of the emitter is determined by current-handling requirements. The base series resistance is reduced by having two base contacts and junction depths are minimized to reduce capacitance. The buried layer, or deep collector, reduces the collector series resistance. High-power transistors are produced by paralleling a number of transistors with interleaving "fingers," as shown in Figure 3.21. This preserves the frequency response of the stripe geometry while increasing the total current-handling capability.

3.1.8 Bipolar Current-Mode Broadband Circuits

Recently there has been strong interest in applying so-called current-mode techniques to electronic circuit design. Considering the signal operating parameter as a current and driving into low-impedance nodes has allowed the development of a wealth of circuits with broadband properties. Many of the following circuit and system concepts date back several years; it is progress in integrated circuit technology that has given a renewed impetus to "practical" current-mode techniques.

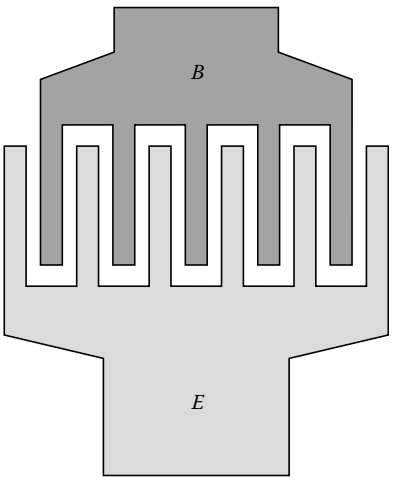


FIGURE 3.21 Transistor layout with interleaving fingers.

The NPN bipolar transistor, for example, is used predominantly in analog IC design because electron mobility is greater than hole mobility in silicon. This means that monolithic structures are typically built on P-type substrates, because vertical NPN transistors are then relatively easy to construct and to isolate from each other by reverse biasing the substrate.

Fabricating a complementary PNP device on a P-type substrate is less readily accomplished. An N type substrate must be created locally and the PNP device placed in this region. Early bipolar processes created PNP devices as lateral transistors and engineers dealt with their inherently poor, low-frequency characteristics by keeping the PNP transistors out of the signal path whenever possible.

However, high-speed analog signal-processing demands symmetrical silicon processes with fully complementary BJTs.

Newer, advanced processes have dielectrically isolated transistors rather than reversed-biased pn junction isolation. These processes are able to create separate transistors, each situated in a local semiconductor region. Then, both PNP and NPN devices are vertical and their performance characteristics are much more closely matched.

Dielectric isolation processes have revolutionized high-speed analog circuit design and have been key in making high-performance current-conveyor and current-feedback op-amp architectures practical. In the following sections, we will briefly review the development of the current-conveyor and current-feedback op-amp.

3.1.8.1 Current Conveyor

The current conveyor is a versatile broadband analog amplifier that is intended to be used with other circuit components to implement many analog signal-processing functions. It is an analog circuit building block in much the same way as a voltage op-amp, but it presents an alternative method of implementing analog systems that traditionally have been based on voltage op-amps. This alternative approach leads to new methods of implementing analog transfer functions, and in many cases the conveyor-based implementation offers improved performance when compared to the voltage op-amp-based implementation in terms of accuracy, bandwidth, and convenience. Circuits based on voltage op-amp are generally easy to design since the behavior of a voltage op-amp can be approximated by a few simple design rules. This is also true for current conveyors, and once the appropriate design rules are understood, the application engineer is able to design conveyor-based circuits just as easily.

The first-generation current conveyor (CCI) was proposed by Smith and Sedra in 1968 [1] and the more versatile second-generation current conveyor (CCII) was introduced by the same two authors in 1970 [2], as an extension of the CCI. The CCII, is without doubt the more valuable and adaptable building block of the two, and we will concentrate mostly on this device. Figure 3.22a shows the voltage–current describing matrix for the CCII, while Figure 3.22b shows the schematic normally used for the CCII with the power supply connections omitted.

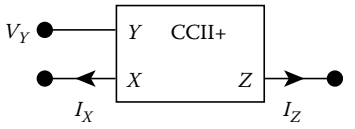
The voltage at the low-impedance input node X follows that at the high-impedance input node Y, while the input current at node X is mirrored or "conveyed" to the high-impedance output node Z. The $\pm$ sign indicates the polarity of the output current with respect to the input current; by convention, a positive sign indicates that both the input and output currents simultaneously flow into or out of the device, thus Figure 3.22b illustrates a CCII+. For the CCI, the input current at node X was reflected to input Y, that is the two inputs had equal currents. In the case of the second-generation conveyor input, Y draws no current, and this second generation, or CCII formulation, has proved to be much more adaptable and versatile than its first-generation predecessor. Because of the combined voltage and current following properties, CCIIs may be used to synthesize a number of analog circuit functions that are not so easily or accurately realizable using voltage op-amps.

Some of these application areas are shown in Figure 3.23. As current-conveyors become more readily available and circuits designers become more familiar with the versatility of this device, it is certain that further ingenious uses will be devised.

The ideal transistor and the current-conveyor. So far a transistor-level realization of the CCII has not been discussed. The current–voltage transfer relationship for the CCII+ is given by

$$\begin{bmatrix} I_Y \\ V_X \\ I_Z \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & \pm 1 & 0 \end{bmatrix} \begin{bmatrix} V_Y \\ I_X \\ V_Z \end{bmatrix}$$

(a)



(b)

FIGURE 3.22 The CCII current conveyor. (a) I–V describing matrix. (b) Schematic.

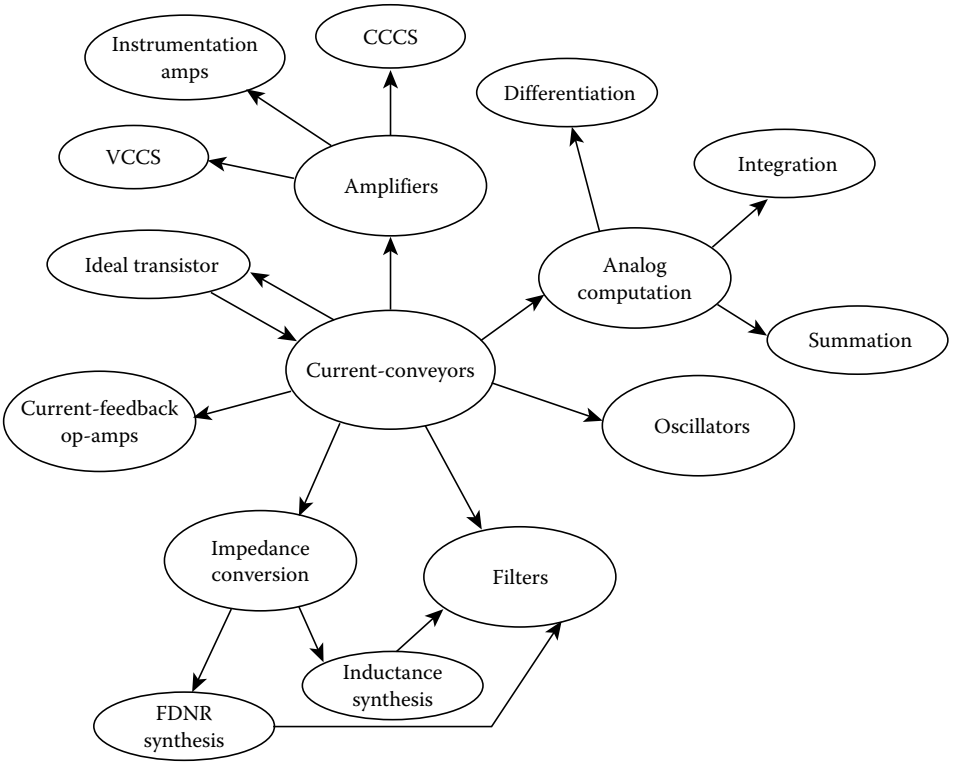


FIGURE 3.23 Current-conveyor applications.

$$V_X = V_Y, \quad I_Y = 0, \quad \text{and} \quad I_Z = I_X \tag{3.23}$$

These equations show that a simple voltage-following action exists between input node Y and output node X, and that there is a simple current-following action between input node X and output node Z. Also, these characteristic equations tell us that the impedance relationship for the ideal current conveyor is

$$Z_{\text{in}Y} = \infty, \quad Z_X = 0, \quad \text{and} \quad Z_{\text{out}Z} = \infty \tag{3.24}$$

Figure 3.24 shows a schematic representation of a CCII—built with a single BJT and on reflection it is clear that the current conveyor is effectively an ideal transistor, with infinite β and infinite g_m.

Driving into the base of a BJT gives almost unity voltage gain from input base to output emitter, with high input impedance and low-output impedance, and driving into the emitter of a BJT gives almost unity current gain from emitter input to collector output, with low input impedance and high output impedance. Drawing the comparison further, the high-input-impedance Y node corresponds to the base (or gate) of a transistor, the low-input-impedance X node corresponds to the emitter (or source) of a transistor, and the high-output-impedance Z node corresponds to the collector (or drain) of a transistor. Clearly, one transistor cannot function alone as a complete current conveyor since an unbiased single transistor at best can only handle unipolar signals and the high-accuracy unity voltage and unity current gain required for a high-performance current conveyor cannot be obtained. However, the generic relationship between the current conveyor and an ideal transistor is valid, and it provides valuable insight into the development and operation of monolithic current conveyors described in the next section.

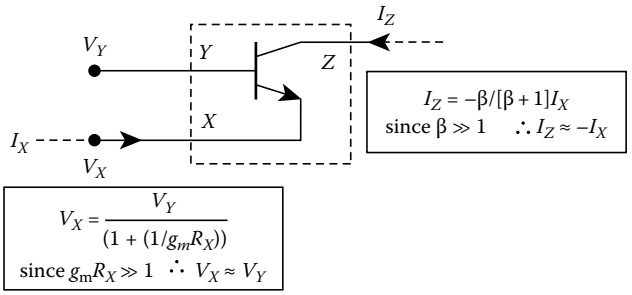


FIGURE 3.24 Single BJT CCII–.

Supply-current sensing. Many of the current-conveyor theories and applications have been tested out in practice using "breadboard" conveyor circuits, due to the lack of availability of a commercial device. Some researchers have built current conveyors from matched transistor arrays, but the most common way of implementing a fairly high-performance current conveyor has been based on the use of supply-current sensing on a voltage op-amp [3,4], as shown in Figure 3.25. The high-resistance op-amp input provides the current-conveyor *Y* node, while the action of negative feedback provides the low-resistance *X* node. Current-mirrors in the op-amp supply leads copy the current at node *X* to node *Z*.

Using this type of architecture, several interesting features soon became apparent. Consider the two examples shown in Figure 3.26. In Figure 3.26b, R_s represents the output resistance of the current source. The open-loop gain of an op-amp can generally be written

$$\frac{V_{out}}{V_{in}} = \frac{A_o}{1 + j(f/f_o)} \tag{3.25}$$

where A_o is the open-loop direct current (dc) gain magnitude and ∂_o is the open-loop -3 dB bandwidth. Since $A_o \gg 1$, the transfer function of the voltage follower of Figure 3.26a can be written as

$$\frac{V_{out}}{V_{in}} \approx \frac{1}{1 + j(f/GB)} \tag{3.26}$$

where $GB = A_o \partial_o$. From Equation 3.26, the -3 dB bandwidth of the closed-loop voltage follower is equal to the open-loop gain-bandwidth product or *GB* of the op-amp. If the op-amp is configured instead to give a closed-loop voltage gain *K*, it is well known that the closed-loop bandwidth correspondingly reduces by the factor *K*.

The transfer function for the current-follower circuit of Figure 3.26b, as shown in Ref. [4], is given by

$$\frac{I_{out}}{I_{in}} \approx \lambda \frac{1 + j(f/GB)}{1 + j(f/kGB)} \tag{3.27}$$

where λ is the current transfer ratio of the current mirrors and $k = (R_s + r_o/A_o)/(R_s + r_o)$, and r_o represents the output resistance of the op-amp. Since $A_o \gg R_s \gg r_o$, then $K \approx 1$, and the pole and zero in Equation 3.27 almost

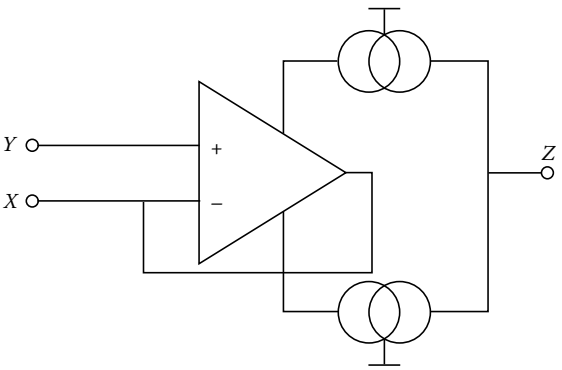


FIGURE 3.25 Supply-current sensing on a voltage op-amp.

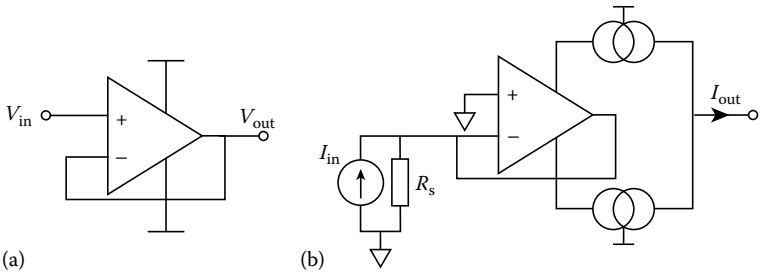


FIGURE 3.26 (a) Voltage follower. (b) Current follower.

cancel. The current-follower circuit thus operates well above the gain-bandwidth product *GB* of the op-amp, and the −3 dB frequency of this circuit will be determined by higher frequency parasitic poles within the current mirrors.

This "extra" bandwidth is achieved because the op-amp is being used with input and output nodes held at virtual ground. The above example is generic in the development of many of the circuits that follow. It demonstrates that reconfiguring a circuit topology to operate with current signals can often result in a superior frequency performance.

First-generation current conveyors. Smith and Sedra's original paper presenting the first-generation CCI current conveyor showed a transistor-level implementation based on discrete devices, shown in Figure 3.27. Assuming that transistors Q_3–Q_5 and resistors R_1–R_3 are matched, then to first order the currents through these matched components will be equal. Transistors Q_1 and Q_2 are thus forced to have equal currents, and equal V_{be}s. Input nodes *X* and *Y* therefore track each other in both voltage and current. In practice, there will be slight differences in the collector currents in the different transistors, due to the finite β of the devices. These differences can be reduced, for example, by using more elaborate current mirrors. The polarity of the output current at node *Z* can be inverted easily by using an additional mirror stage, and the entire circuit can also be inverted by replacing NPN transistors with PNPs, and vice versa. Connecting two complementary current conveyors, as shown in Figure 3.28, results in a class *AB* circuit capable of bipolar operation. Note that in practice this circuit may require additional components to guarantee start-up.

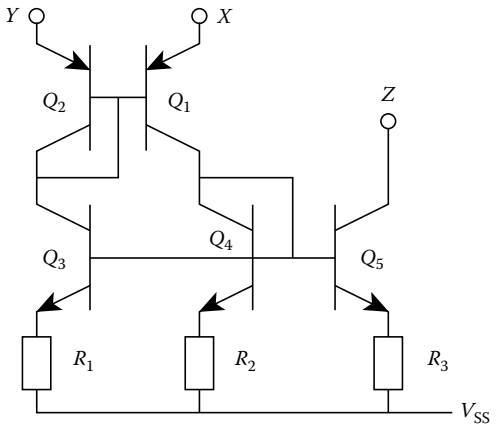


FIGURE 3.27 First-generation current conveyor.

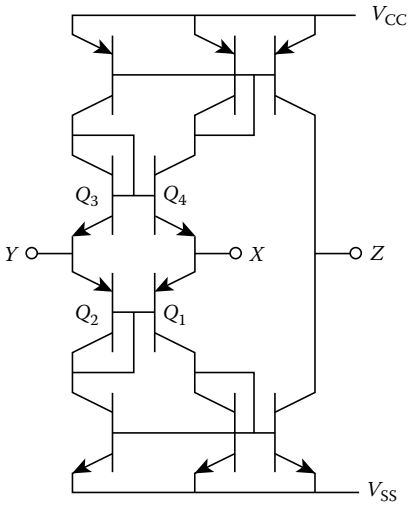


FIGURE 3.28 Class *AB* current conveyor.

An integrated current conveyor based on the architecture shown in Figure 3.27 is commercially available as the PA630 [5], and the basic topology of this device is shown in Figure 3.29. An NPN Wilson mirror (Q_1–Q_3) and a PNP Wilson mirror (Q_4–Q_6) are used to provide the current and voltage following properties between inputs X and Y, similar to the circuit of Figure 3.27. Taking a second output from the PNP current mirror to provide the Z output would destroy the base-current compensation scheme of the Wilson mirror. Therefore, a second NPN Wilson mirror (Q_7–Q_9) is used to perform a current-splitting action and so the combined emitter current of Q_7 and Q_8 is divided in two, with one half being shunted via Q_9 to the supply rail, and the other half driving an output PNP Wilson mirror (Q_{10}–Q_{12}). This results in an output current at node Z that to first order is virtually equal to that at the X and Y inputs. Q_{13} is included to ensure that the device always starts up when turned on. The complete architecture of the PA630 CCI also includes frequency compensation to ensure stability, and modified output current mirrors that use the ''wasted'' collector current of Q_9 to effectively double the output resistance at node Z. A full description of the architecture and operation of this device can be found in Ref. [6].

The current-conveyor architecture shown in Figure 3.29 includes both NPN and PNP transistors in the signal path, and thus the bandwidth and current-handling capability of this device will be poor if only lateral PNPs are available. The development of complementary bipolar processes, with vertical PNP as well as NPN transistors, has made possible the implementation of high-performance integrated circuit current conveyors.

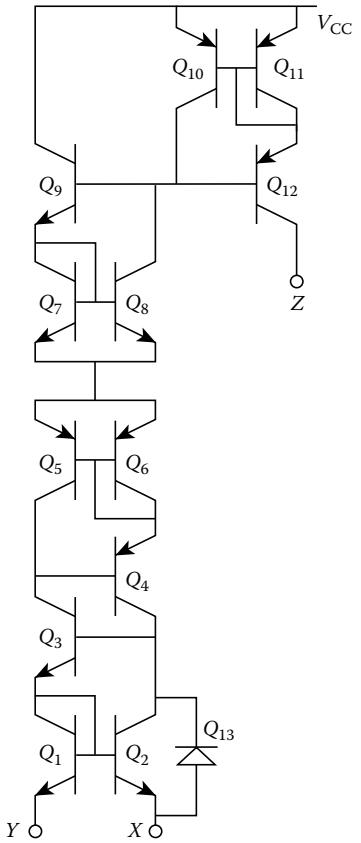


FIGURE 3.29 Simplified PA630 current conveyor.

Second-generation current conveyors. A CCII can also be simply implemented on a complementary bipolar process, by replacing the diode at the CCI Y input with a transistor, and taking the input from the high resistance base terminal, as shown in Figure 3.30a. This can be extended to a class AB version, as shown in Figure 3.30b. Referring to Figure 3.30b, transistors Q_1–Q_4 act as a voltage buffer that transfers the voltage at node Y to node X. The current source and sink ($I_{B1} = I_{B2} = I_B$) provide the quiescent bias current for these input transistors. Any input current (I_x) at node X is split between Q_2 and Q_3, and is copied by current mirrors CM_1 and CM_2 to the output node Z. This CCII architecture forms the basis of the commercially available CCII01 current conveyor [7]. As we shall see later, it is also used as the basic input stage of the current-feedback op-amp, which has emerged as a high-speed alternative to the more conventional voltage op-amp [8].

The simple CCII architecture of Figure 3.30b will clearly exhibit a quiescent voltage offset between nodes X and Y due to the mismatch between the V_{be}s of the NPN and PNP transistors Q_1/Q_2 and Q_3/Q_4, as

$$\begin{aligned} V_Y - V_X &= V_{BE}(p) - V_{BE}(n) \\ &= V_T \ln\left(I_{sp}/I_{sn}\right) \end{aligned} \tag{3.28}$$

where I_{sp} and I_{sn} are the reverse saturation currents of the PNP and NPN transistors, respectively, and V_T is the thermal voltage. This process-dependent voltage offset can be reduced by including additional matching diodes in the input stage, as shown in Figure 3.31. Referring to this diagram,

$$\begin{aligned} V_Y - V_X &= V_{BE}(Q_1) + V_{D_2} - V_{BE}(Q_2) - V_{D_1} \\ V_Y - V_X &= [V_{BE}(Q_1) - V_{D_1}] - [V_{BE}(Q_2) - V_{D_2}] \end{aligned} \tag{3.29}$$

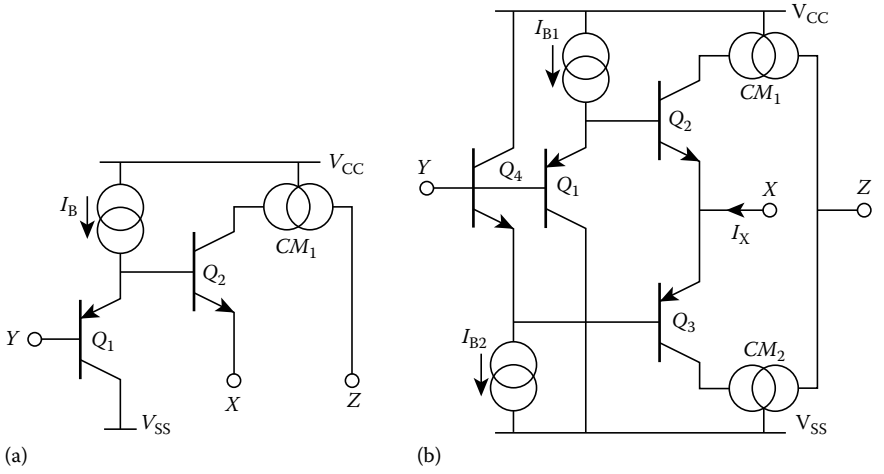


FIGURE 3.30 (a) Class *A* CCII. (b) Class *AB* CCII.

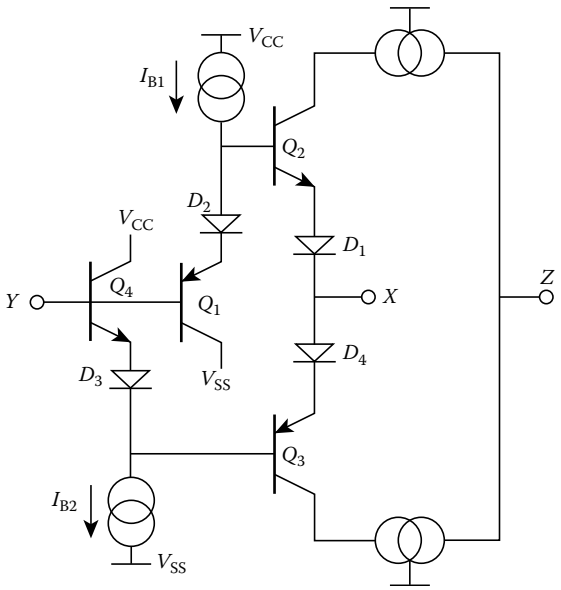


FIGURE 3.31 CCII with input matching diodes.

Inclusion of these diodes clearly reduces the quiescent input voltage offset, provided that D_1 is matched to Q_1, D_2 is matched to Q_2, etc. However, the addition of diodes D_1 and D_2 has several disadvantages. First, the input voltage dynamic range of the circuit will be reduced by the forward voltage across the additional diode. Second, the small-signal input resistance seen looking into node X will be double that for the basic architecture given in Figure 3.30b. This nonzero input resistance at node X (R_x) will compromise the performance of the current conveyor, especially in applications where a nonzero input voltage is applied at node Y. The effect of the small-signal input resistance R_x is to produce a signal-dependent voltage offset V_d between nodes X and Y, where

$$V_{\dot{a}} = R_x I_x \tag{3.30}$$

Since the value of R_x is determined by the small-signal resistance $(r_{e2} + r_{d2})$ in parallel with $(r_{e3} + r_{d3})$, its value could be reduced by increasing the value of the quiescent bias current I_B. However, an increase in bias current will lead to an increase in the total power consumption, as well as a possible increase in offsets, and so is certainly not an ideal solution. Further techniques for CCII implementation are discussed in Ref. [14].

The previous conveyor is typical of commercial conveyor architectures [7], which are generally built on a high-speed dielectric isolation (fully complementary) bipolar process. Such devices feature an equivalent slew rate of some 2000 V/μs and a bandwidth of around 100 MHz.

Until high-performance current conveyors are widely available, these devices will continue to be used in research laboratories rather than in the applications arena. Process technologies and design techniques

have now advanced to the stage where the implementation of an integrated current conveyor is both desirable and viable, and a whole host of applications are waiting for its arrival.

3.1.8.2 Current-Feedback Operations Amplifier

In this section, the design and development of a high-gain wide-bandwidth transimpedance or current-feedback operational amplifier is considered. The design of conventional operational amplifiers has remained relatively unchanged since the introduction of the commercial operational amplifier in 1965. Recently, a new amplifier architecture, called a current-feedback operational amplifier, has been introduced. This amplifier architecture is basically a transimpedance amplifier, or a current-controlled voltage source, while the classical voltage-feedback operational amplifier is a voltage-controlled voltage source.

The current-feedback operational amplifier has two major advantages, compared to its voltage-feedback counterpart. First, the closed-loop bandwidth of the current-feedback amplifier is larger than that of classical voltage-feedback design for comparable open-loop voltage gain. Second, the current-feedback operational amplifier is able to provide a constant closed-loop bandwidth for closed-loop voltage gains up to about 10. A further advantage of the current-feedback architecture is an almost unlimited slew rate due to the class-*AB* input drive, which does not limit the amount of current available to charge up the compensation capacitor as is the case in the conventional voltage-feedback op-amp. This high-speed performance of the current-feedback operational amplifier is extremely useful for analog signal-processing applications within video and telecommunication systems.

The generic relationship between the CCII+ and the current-feedback op-amp is extremely close and several of the features offered by the CCII are also present in the current-feedback op-amp. The basic structure of the current-feedback op-amp is essentially that of a CCII+ with the Z node connected directly to an output voltage follower, as shown in Figure 3.32. Any current flowing into the low-impedance inverting input is conveyed to the gain node (Z_T), and the resulting voltage is buffered to the output. Z_T is thus the open-loop transimpedance gain of the current-feedback op-amp, which in practice is equal to the parallel combination of the CCII+ output impedance, the voltage buffer input impedance and any additional compensation capacitance at the gain node. Generally, in current-feedback op-amps, the gain node is not connected to an external pin, and so the Z node of the CCII+ cannot be accessed.

Current-feedback op-amp architecture. In the following sections, we review the basic theory and design of the current-feedback op-amp and will identify the important features and mechanisms that result in broadband performance. We will begin by reviewing the voltage-feedback op-amp and comparing it with the current-feedback op-amp in order to see the differences clearly.

A schematic of the classical voltage-feedback op-amp comprising a long-tail pair input stage is shown in Figure 3.33a, which contrasts a typical current-feedback architecture, which is shown in Figure 3.33b. In both circuits, current mirrors are represented by two interlocking circles with an arrow denoting the input side of the mirror.

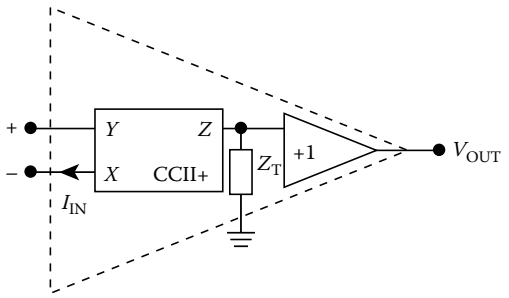


FIGURE 3.32 Current-feedback op-amp structure.

The current-feedback op-amp of Figure 3.33b shows that the noniverting input is a high-impedance input that is buffered to a low-impedance inverting terminal via a class *AB* complementary common-collector stage (Q_1, Q_2, D_1, D_2). Note that this classical input buffer architecture is used here for simplicity. In practice, a higher performance topology such as that described in Figure 3.31 would more likely be employed. The noninverting input is a voltage input; this voltage is then buffered to the inverting low-impedance current input to which feedback is applied. In contrast, both the noninverting and

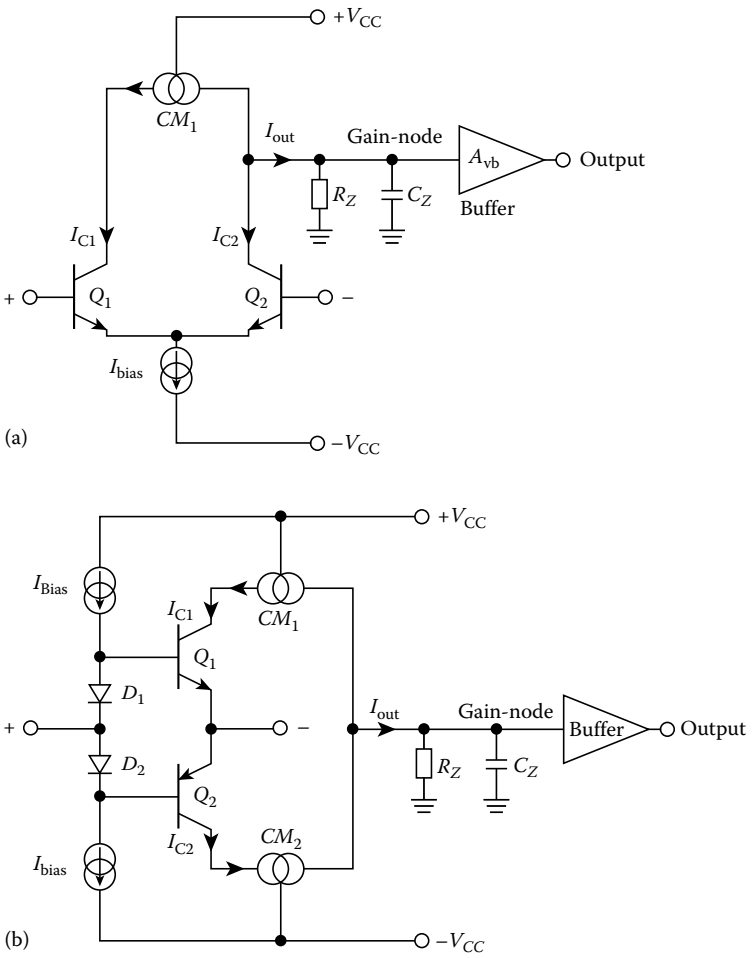


FIGURE 3.33 (a) Simplified classic voltage-feedback op-amp architecture. (b) Typical current-feedback op-amp architecture.

inverting input of the voltage-feedback op-amp are high-impedance voltage inputs at the bases of transistors Q_1 and Q_2.

In both architectures, the collector currents of Q_1 and Q_2 are transferred by the current mirror to a high-impedance node represented by resistance R_Z and capacitance C_Z. This voltage is then transferred to the output by voltage buffers that have a voltage gain A_{vb}, providing the necessary low-output impedance for current driving. In the case of the current-feedback op-amp, the output buffer is usually the same topology as the input buffer stage shown in the Figure 3.33b, but with slightly higher output current bias levels and larger output devices to provide an adequate output drive capability. Ideally, the bias currents I_{CQ1} and I_{CQ2} will be canceled at the gain node giving zero offset current.

Differential-mode operation of the current-feedback op-amp. A schematic diagram of the current-feedback op-amp with a differential input voltage applied at the noninverting and inverting input is shown in Figure 3.34.

The positive input voltage is applied to the base of transistor Q_1 (NPN) via D_1, and the negative input voltage is applied to the emitter of Q_1, causing the V_{BE} of Q_1 to increase and the V_{BE} of Q_2 to

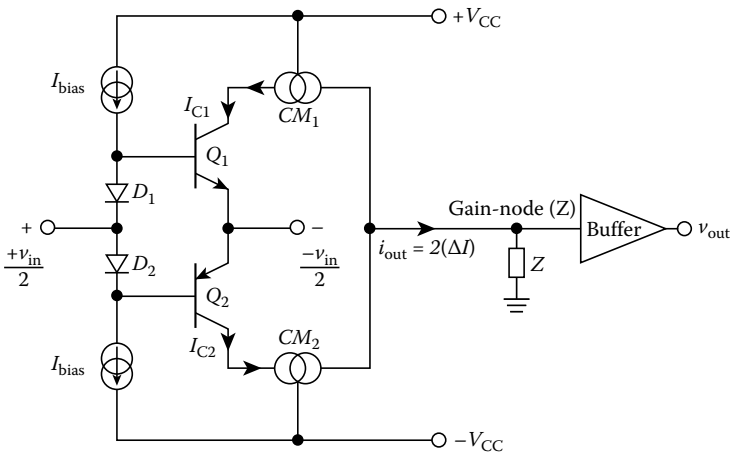


FIGURE 3.34 Current-feedback op-amp with differential input voltage applied.

reduce. I_{C1} will therefore increase by an amount ΔI and so I_{C2} will decrease by the same amount $-\Delta I$. A net current of $2\Delta I$ is therefore sourced out of the high-impedance node (Z) giving rise to a positive voltage ($2\Delta IZ$). This voltage is then buffered to the output.

With negative feedback applied around the current-feedback op-amp, the low-impedance inverting input will sense the current "feedback" from the output via the feedback network. This feedback current flowing into the inverting input is given by

$$i_{in-} = I_{C2} - I_{C1} \tag{3.31}$$

The difference between the collector current I_{C1} and I_{C2}, i_{in-}, will thus be driven into gain node Z, giving rise to the output voltage

$$V_{out} = Zi_{in-} \tag{3.32}$$

It is clear that the output voltage is dependent on the current that flows into the inverting input, hence the amplifier has a high open-loop transimpedance gain Z.

Closed-loop noninverting operation of the current-feedback op-amp. A schematic diagram of the current-feedback op-amp connected with negative feedback as a noninverting amplifier is shown in Figure 3.35. For a positive input voltage v_{in}, the output voltage v_{out} will swing in the positive direction and the inverting input current i_{in-} will flow out:

$$i_{in-} = \frac{v_{in-}}{R_1} - \frac{(v_{out} - v_{in-})}{R_2} \tag{3.33}$$

The input stage is simply a voltage follower and so ideally, $v_{in+} = v_{in-} = v_{in}$. Because $v_{out} = Zi_{in-}$, then substituting for v_{in-} and i_{in-} in Equation 3.33 yields

$$\frac{v_{out}}{Z} = \frac{v_{in}}{R_1} - \frac{(v_{out} - v_{in})}{R_2} \tag{3.34}$$

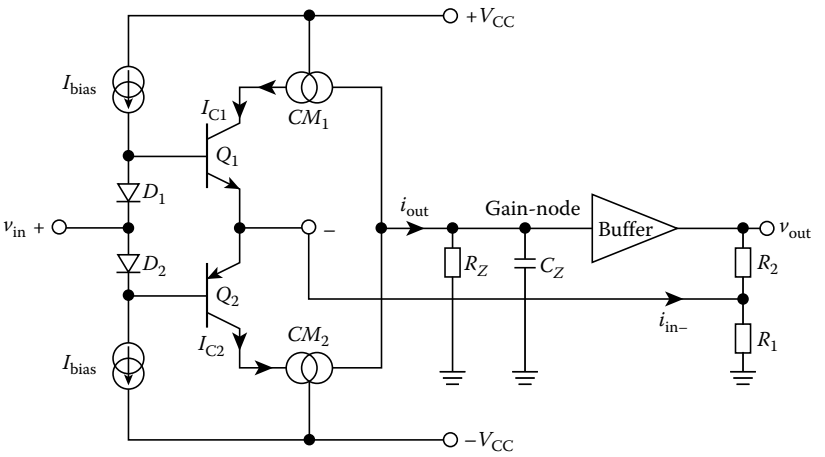


FIGURE 3.35 Noninverting current-feedback op-amp.

rearranging for v_{out}/v_{in}

$$v_{out}\left(\frac{1}{R_2}+\frac{1}{Z}\right)=v_{in}\left(\frac{1}{R_1}+\frac{1}{R_2}\right) \tag{3.35}$$

$$\frac{v_{out}}{v_{in}}=\left(1+\frac{R_2}{R_1}\right)\left(\frac{1}{1+(R_2/Z)}\right) \tag{3.36}$$

This result shows that the closed-loop noninverting gain of the current-feedback op-amp is similar to that of a classical voltage-feedback op-amp. From Equation 3.36, the open-loop transimpedance gain Z must be as large as possible to give good closed-loop gain accuracy. Since v_{out}/Z represents the error current i_{in-}, then maximizing the Z term will minimize the inverting error current. Note that at this stage it is only the R_2 term in the denominator of the second term in Equation 3.36 that sets the bandwidth of the amplifier; the gain-setting resistor R_1 has no effect on the closed-loop bandwidth.

Closed-loop inverting operation of current-feedback op-amp. A current-feedback op-amp connected as an inverting amplifier is shown in Figure 3.36. The low-impedance inverting input samples the input current and drives the output until the voltage at its terminal is at a virtual ground because of negative feedback. Ideally the closed-loop gain is given by

$$A_{CL}=-\frac{R_2}{R_1} \tag{3.37}$$

From Figure 3.36, application of Kirchhoff's current law to the current i_1, i_{in-}, and i_2 gives

$$i_{in-}+i_2=i_1$$
$$i_{in-}-\frac{v_{out}}{R_2}=\frac{v_{in}}{R_1}$$

because $v_{out}/Z=-i_{in-}$, then

$$-\frac{v_{out}}{Z}-\frac{v_{out}}{R_2}=\frac{v_{in}}{R_1}$$

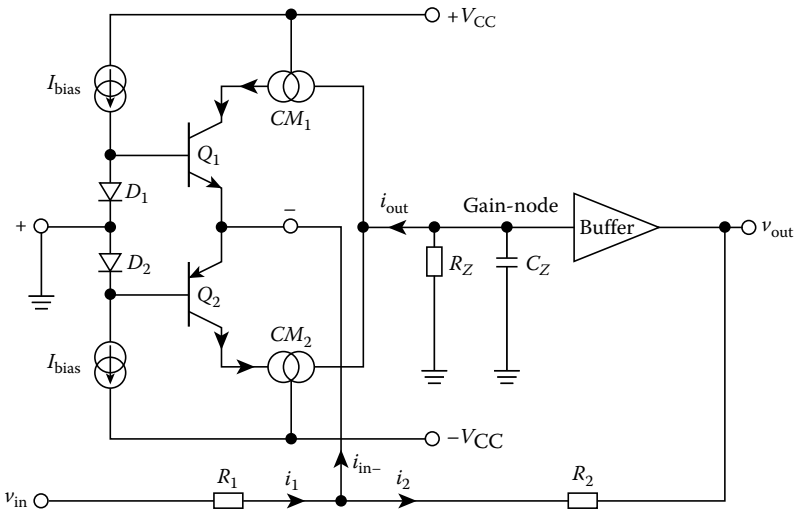


FIGURE 3.36 Inverting current-feedback op-amp amplifier.

which can be rearranged as

$$\frac{v_{\text{out}}}{v_{\text{in}}} = -\frac{R_2}{R_1}\left(\frac{1}{1+\frac{R_2}{Z}}\right) \tag{3.38}$$

Again, the high-Z term is required to provide good closed-loop gain accuracy.

More detailed analysis of the current-feedback op-amp. A simplified macromodel of the current-feedback architecture configured as a noninverting amplifier is shown in Figure 3.37. The input stage is represented by a semi-ideal voltage buffer to the inverting input. The output resistance of the input stage buffer R_{inv} is included since it has a significant effect on the bandwidth of the amplifier, as will be shown later. The current that flows out from the inverting terminal i_3 is transferred to the gain node, which is represented by R_Z and C_Z, via a current mirror that has a current gain K. The voltage at the gain node is transferred to the output in the usual way by a voltage buffer, with voltage gain A_{vb}. The net transfer function is given by

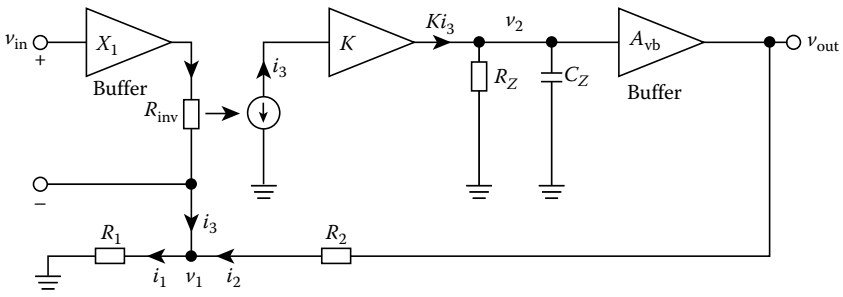


FIGURE 3.37 Inverting amplifier with current-feedback op-amp macromodel.

$$\frac{v_{out}}{v_{in}} = \frac{1 + \dfrac{R_2}{R_1}}{1 + j\omega C_Z \left[\dfrac{R_{inv}\left(1 + \dfrac{R_2}{R_1}\right) + R_2}{A_{vb}K} \right]} \tag{3.39}$$

Hence, the pole frequency is also given by

$$f_{-3dB} = \frac{A_{vb}K}{2\pi C_Z \left[R_{inv}\left(1 + \dfrac{R_2}{R_1}\right) + R_2 \right]} \tag{3.40}$$

(A full derivation of this transfer function is given in Appendix A.)

To compare this result to the classical voltage-mode op-amp architecture, a simplified schematic diagram of the voltage-feedback op-amp configured as a noninverting amplifier is shown in Figure 3.38.

Again from a full analysis, given in Appendix B, the transfer function obtained is

$$\frac{v_{out}}{v_{in}} = \frac{1 + \dfrac{R_2}{R_1}}{1 + j\omega \left[\dfrac{R_z C_Z}{1 + \dfrac{g_m A_{vb} R_Z}{\left(1 + \dfrac{R_2}{R_1}\right)}} \right]} \tag{3.41}$$

The pole frequency is given by

$$f_{-3dB} = \frac{1 + \dfrac{g_m A_{vb} R_Z}{\left(1 + \dfrac{R_2}{R_1}\right)}}{2\pi R_Z C_Z} \tag{3.42}$$

Pole frequency comparison. If one compares the closed-loop pole frequency Equations 3.40 and 3.42 for the current-feedback and voltage-feedback op-amp, respectively, it is clear that the bandwidth of the voltage-feedback op-amp is dependent on the closed-loop gain $(1 + R_2/R_1)$ resulting in the well-known constant gain-bandwidth product $f_{max} = (A_v)_{CL} f_T$. This means that an increase in the closed-loop gain results in a decrease in the bandwidth by the same factor as illustrated in Figure 3.39. In contrast, the pole

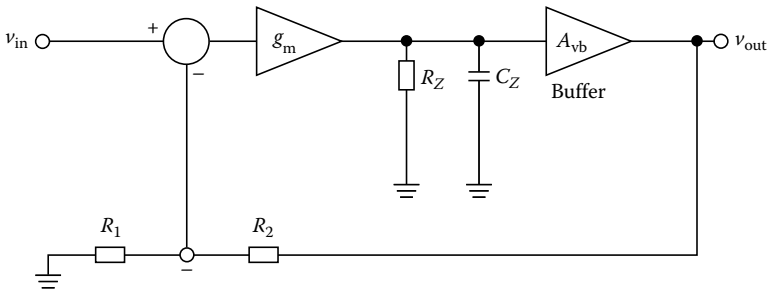


FIGURE 3.38 Noninverting amplifier with voltage-feedback op-amp macromodel.

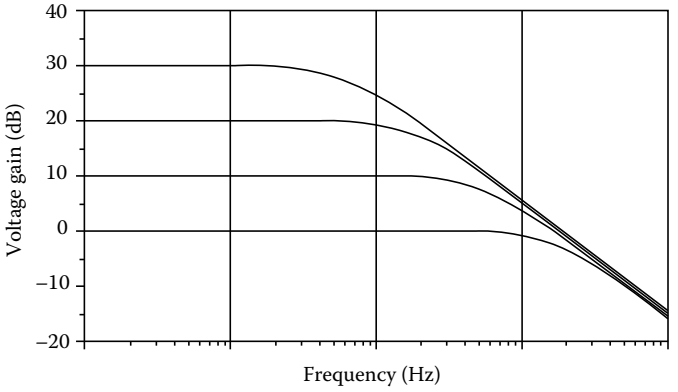


FIGURE 3.39 Frequency response of voltage-feedback op-amp amplifier for various closed-loop gains.

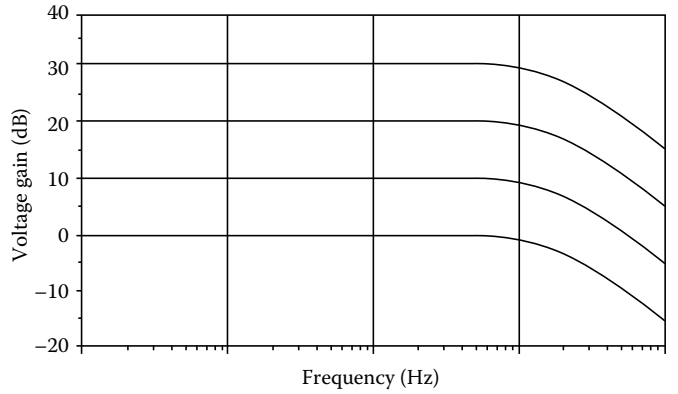


FIGURE 3.40 Frequency response of current-feedback op-amp amplifier for various closed-loop gains.

frequency of the current-feedback op-amp is directly dependent on R_2 and can be set almost independently of the closed-loop gain. Thus, the closed-loop bandwidth is almost independent of closed-loop gain as shown in Figure 3.40, assuming that R_{inv} is close to zero. Intuitively, this is the case since the feedback error current that is set by the feedback resistor R_2 is the current available to charge up the compensation capacitor. However, if one considers Equation 3.40 in some detail it can be seen that for high closed-loop gains and a nonzero R_{inv}, then the R_{inv} term starts to dictate and so the bandwidth will become more dependent on the closed-loop gain.

Slow rate of the current-feedback op-amp. As mentioned earlier, one other advantage of the current-feedback op-amp over the classical voltage-feedback op-amp is the high slew rate performance. For the classical long-tail, or emitter-coupled pair input stage shown in Figure 3.41, the maximum current available to charge up the compensation capacitor C_Z at the gain node is I_{bias}, and this occurs when Q_1 or Q_2 is driven fully on. The resulting transconductance plot shown in Figure 3.42 limits the slew rate of the amplifier.

In contrast, the slew rate of the current-feedback op-amp is virtually infinite, as can be seen from the input stage schematic shown in Figure 3.43. Referring to Figure 3.43, a change in the input voltage ΔV_{in} at $V(+)$ will be copied by the input buffer to $V(-)$. When connected as noninverting amplifier, the current through R_1 will change by $\Delta V_{in}/R_1$, while the current through R_2 will change by $\Delta V_{in}/R_2$, since

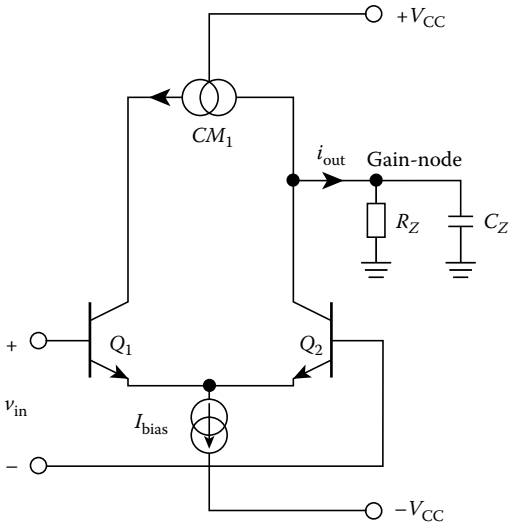


FIGURE 3.41 Long-tail pair input stage.

the output voltage at this point remains stationary. The total change in current through R_1 and R_2 must be supplied by the internal input buffer, and will be $\Delta I(-) = \Delta V_{in}((R_2 + R_1)/(R_2 \times R_1))$. This large input error current causes a rapid change in the output voltage, until V_{out} is again at the value required to balance the circuit once more, and reduce $I(-)$ to zero. The larger the input voltage slew rate, the larger the change in input error current, and thus the faster the output voltage slew rate. Current-feedback op-amps theoretically have no slew-rate limit. A typical current-feedback op-amp will exhibit a slew rate of between 500 and 2000 V/μS.

An analysis of this input stage (see Appendix C) shows that the transconductance follows a sinh(x) type function, as shown in Figure 3.44. In theory, this characteristic provides nearly unlimited slew-rate capability [9]. However, in practice a maximum slew rate will be limited by the maximum current drive into the gain node, which depends on the power dissipation of the circuit, the ability of power supply to deliver sufficient current, and the current-handling capability of the current mirrors.

Wideband and high-gain current-feedback op-amp. Previously, we have shown that the bandwidth of the current-feedback op-amp is almost independent of the closed-loop gain setting. Therefore, the closed-loop gain-bandwidth GB increases linearly with the closed-loop gain. However, the bandwidth of the practical current-feedback op-amp starts decreasing with high gain as a result of the finite inverting-input impedance [10], as shown by Equation 3.40. This is because for high gain, $R_{inv}(1 + R_2/R_1) > R_2$, and so the $R_{inv}(1 + R_2/R_1)$ term dominates the expression for closed-loop bandwidth, resulting in a direct conflict between gain and bandwidth.

At low gains when $R_2 > R_{inv}(1 + R_2/R_1)$, the closed-loop pole frequency is determined only by the compensation capacitor and the feedback resistor R_2. Thus, the absolute value of the feedback resistor R_2 is important, unlike the case of the voltage-feedback op-amp. Usually, the manufacturer species a minimum value of R_2 that will maximize bandwidth but still ensure stability. Note that because of the

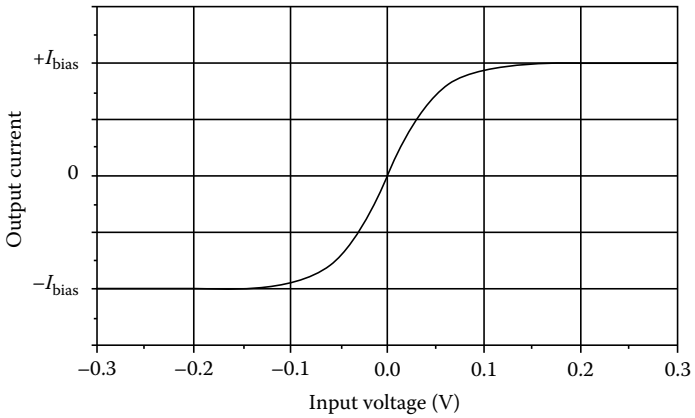


FIGURE 3.42 Long-tail pair input transconductance.

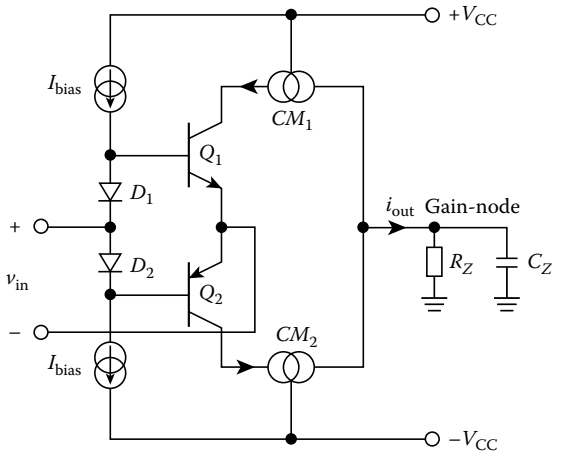


FIGURE 3.43 Current-feedback op-amp input stage.

inherent architecture a very high bandwidth can be achieved with the current-feedback design for a given value of R_2.

In practice, for gains higher than about 10, the $R_{inv}(1 + R_2/R_1)$ term in Equation 3.40 becomes dominant and the amplifier moves toward constant gain-bandwidth product behavior. The *GB* can be increased by reducing R_2 [11] but this will compromise stability and/or bandwidth, or alternatively, C_Z can be reduced. The latter option is limited since the minimum value of C_Z is determined by the device parameters and layout parasitics. Two possible ways of improving the high-gain constant bandwidth capability of the current-feedback op-amp can be seen by inspection of Equation 3.40. Either the K factor, which represents current gain in the current mirrors at the Z-node can be increased from unity to increase the bandwidth as it rolls off with high gain, or the inverting input impedance of the amplifier should be reduced toward zero. In the following section we consider the design of a suitable broadband variable-gain current-mirror circuit with a possible application being to improving the maximum bandwidth capability of current-feedback op-amps.

Basic current mirror. A typical current-feedback op-amp circuit is shown in Figure 3.45. It includes a complementary common-collector input stage (Q_1–Q_4) and a similar output buffer (Q_5–Q_8), with linking cascode current mirrors setting the Z-node impedance (Q_{12}–Q_{14}, Q_9–Q_{11}). The cascoded mirror provides unity current gain. Any attempt to increase the current gain via emitter degeneration usually results in much poorer current-mirror bandwidth. Consider now the development of a suitable broadband, variable gain current mirror.

A schematic diagram of a simple Widlar current mirror and its small-signal equivalent circuit are shown in Figures 3.46 and 3.47, respectively. For simplicity, we will assume that the impedance of the diode-connected transistor Q_1 is resistive and equal to R_D. The dc transfer function of the mirror is derived in Appendix D and is given by

$$\frac{I_{out}}{I_{in}} = \frac{\beta}{\beta + 2} \tag{3.43}$$

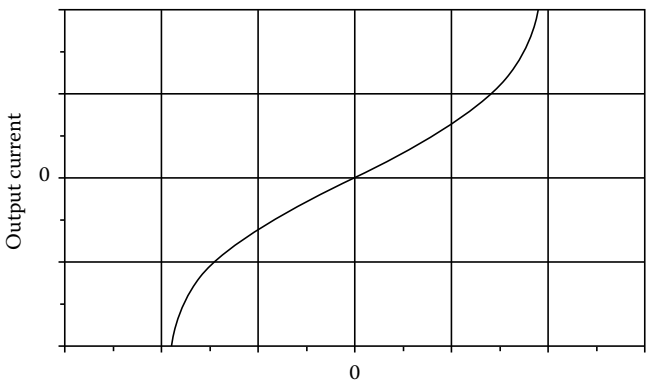


FIGURE 3.44 Input-stage transconductance of the current-feedback op-amp.

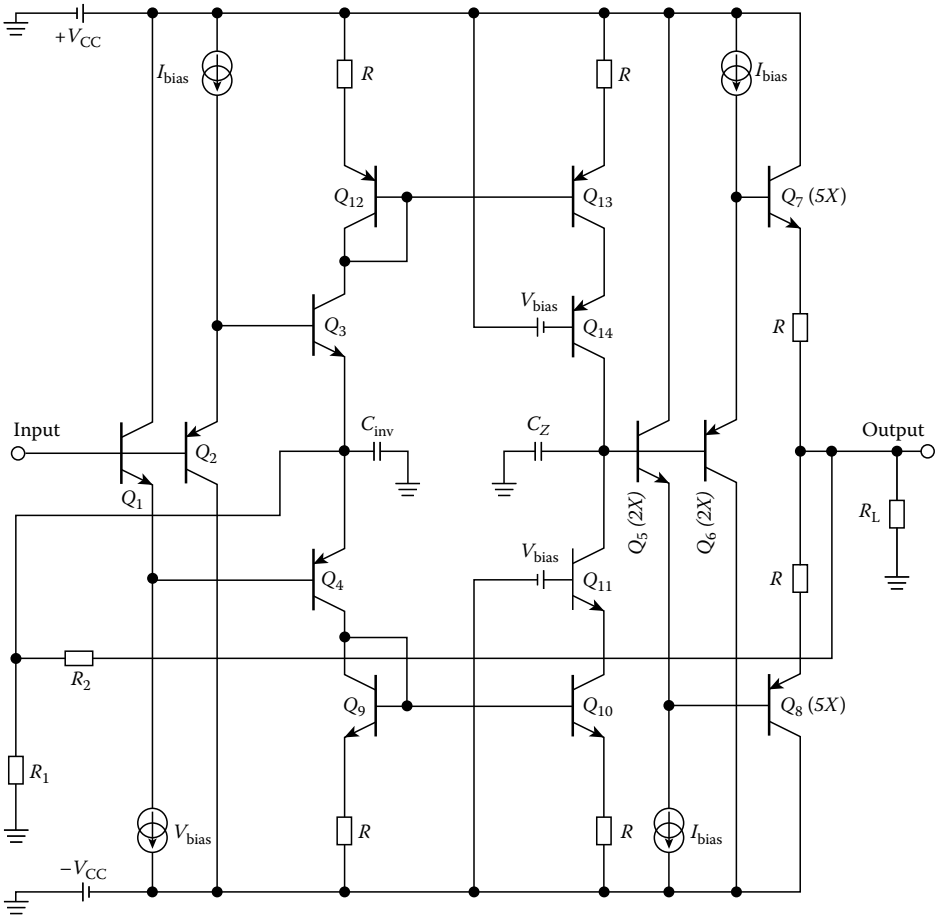


FIGURE 3.45 Transistor-level schematic of a typical current-feedback op-amp. $x =$ Unit transistor area.

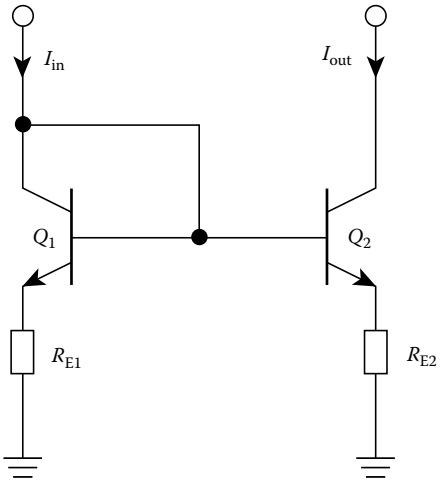


FIGURE 3.46 Simple Widlar current mirror with emitter degeneration.

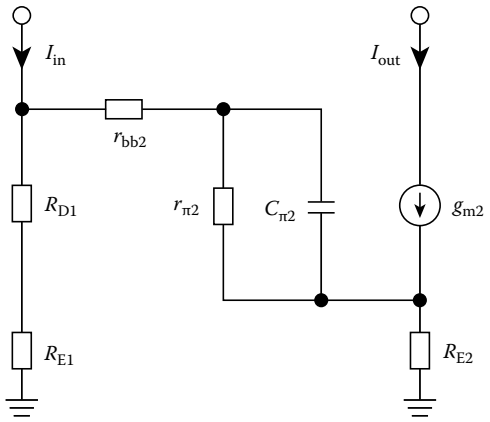


FIGURE 3.47 Small-signal equivalent circuit of Figure 3.46 current mirror.

and the −3 dB bandwidth is given by

$$f_{-3\text{ dB}} = \frac{1}{2\pi C_\pi \left\{ \dfrac{r_{\pi 2}(r_{\text{bb2}} + R_D)}{r_\pi + r_{\text{bb2}} + R_D} \right\}} \tag{3.44}$$

In order to increase the current gain it is usual to insert an emitter-degeneration resistor R_{E1} in the emitter of Q_1. The dc transfer function, derived in the Appendix E, is then

$$I_{\text{in}} R_{\text{E1}} = V_{\text{T}} \ln \frac{I_{\text{out}}}{I_{\text{in}}} \tag{3.45}$$

and the ac small-signal current gain is given by

$$\frac{i_{\text{out}}}{i_{\text{in}}} = (R_{\text{E1}} + R_{\text{D1}}) g_{\text{m2}} \tag{3.46}$$

where

$$R_{\text{D1}} = \frac{I_{\text{in}}}{\frac{KT}{q}} \tag{3.47}$$

The −3-dB bandwidth now becomes

$$f_{-3\text{dB}} = \frac{1}{2\pi C_{\pi 2} \left\{ \dfrac{r_{\pi 2}(r_{\text{bb2}} + R_{\text{D1}} + R_{\text{E1}})}{r_{\pi 2} + r_{\text{bb2}} + R_{\text{D1}} + R_{\text{E1}}} \right\}} \tag{3.48}$$

It can be seen that increasing R_{E1} to increase the gain results in a reduction in the mirror bandwidth. The method of increasing the area of Q_2 to increase the current gain is not advantageous because the capacitance $C_{\pi 2}$ increases simultaneously, and so again, the bandwidth performance is compromised. We can conclude that this approach, though apparently well founded, is flawed in practice.

Improved broadband current mirror. A current mirror with current gain is shown in Figure 3.48 and the small-signal equivalent circuit is shown in Figure 3.49. In this current mirror Q_1 and Q_2 are connected as diodes in series with R_{E1}. Q_3 is connected as a voltage buffer with the bias current source I_{EQ3}. Q_4 is the output transistor with degeneration resistor R_{E4} for current gain setting. The basic idea is to introduce the CC Q_3 to buffer the output from the input and hence isolate gain setting resistor R_{E4} from the bandwidth determining capacitance of the input. The dc transfer function is given by

$$I_{\text{in}} R_{\text{E1}} - I_{\text{out}} R_{\text{E4}} + V_{\text{T}} \ln \frac{I_{\text{in}}^2}{I_{\text{CQ3}} I_{\text{out}}} = 0 \tag{3.49}$$

and the ac small-signal current gain is given by

$$\frac{i_{\text{out}}}{i_{\text{in}}} = \frac{(R_{\text{E1}} + R_{\text{D1}} + R_{\text{D2}}) g_{\text{m4}}}{1 + g_{\text{m4}} R_{\text{E4}}} \tag{3.50}$$

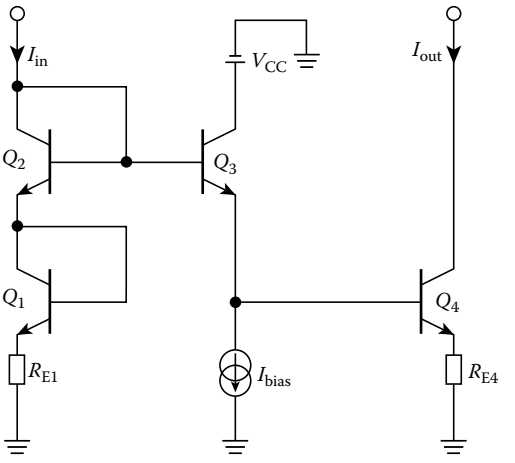


FIGURE 3.48 Improved current mirror with current gain.

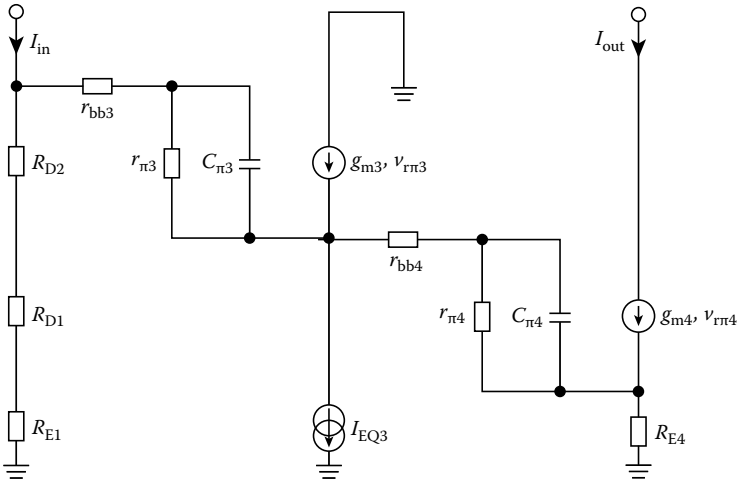


FIGURE 3.49 Equivalent circuit of improved current mirror with current gain.

and the −3 dB bandwidth now becomes

$$f_{-3\text{ dB}} = \frac{1}{2\pi C_{\pi 4}\left(\dfrac{r_{\pi 4}R_x}{r_{\pi 4}+R_x}\right)} \tag{3.51}$$

where

$$R_x = r_{bb4} + \frac{r_{\pi 3} + r_{bb3} + R_{D1} + R_{D2} + R_{E1}}{\beta_3} \tag{3.52}$$

It can be seen clearly that the dominant pole Equation 3.51 of the current mirror with current gain is now only slightly decreased when we increase the current gain by increasing R_{E1}. However, the nondominant pole at the input node is increased, and this will marginally effect the resultant overall stability performance if employed in a current-feedback op-amp. This current mirror with current gain has been employed successfully in current-feedback op-amp design for increased gain-bandwidth capability [12].

Phase linearity. The internal signal path in a current-feedback op-amp is very linear due largely to the symmetrical architecture. Consequently, these devices have a very linear phase response. Furthermore, all the frequency components of a signal are delayed by the same amount when passing through the amplifier, and so the waveform is reproduced accurately at the output. Current-feedback op-amps typically exhibit differential phase error of around ±1° at frequencies of approximately half the bandwidth.

Choosing the value of R_2. From Equation 3.40, we can see that for a fixed value of C_z, a smaller feedback resistor R_2 will give a higher closed-loop bandwidth. It might be expected that the maximum bandwidth would be obtained with the minimum feedback resistance; that is, with $R_2 = 0$. In practice, current-feedback op-amps are generally unstable when their feedback resistance is reduced below a particular value. The reason for this is that the dominant closed-loop pole at frequency of $f \approx 1/2\pi C_z R_2$ must be significantly lower than any nondominant parasitic pole frequency within the op-amp, so that a reasonable gain and phase margin is maintained. If the value of R_2 is reduced, then this dominant pole will move upward in frequency toward the parasitic poles, reducing the gain and phase margin, and eventually leading to instability. Obviously, the "correct" value for R_2 will depend on the internal value of C_z and the location of any parasitic poles within the device. These are the sort of parameters

that are known to the manufacturer, but are generally not listed in a data sheet. Therefore, the manufacturer of a particular device will generally recommend a value of R_2 that guarantees stability, while maintaining a reasonably wide bandwidth. Reducing R_2 below this recommended or optimum value will tend to lead to peaking and instability, while increasing R_2 above the optimum value will reduce the closed-loop bandwidth. If band limiting is required, then a larger value of R_2 than the optimum can be chosen to limit the bandwidth as required.

Since a current-feedback op-amp requires a minimum value of R_2 to guarantee stability, these devices cannot be used with purely capacitive feedback because the reactance of a capacitor reduces at high frequencies. This means that the conventional voltage op-amp integrator cannot be implemented using a current-feedback op-amp.

Practical considerations for broadband designs.

1. *Ground planes.* The purpose of a ground plane is to provide a low-impedance path for currents flowing to ground, since any series impedance in the ground connections will mean that not all ground nodes are at the same potential. In addition, the inductance of a printed circuit track is approximately inversely proportional to the track width, and so the use of thin tracks can result in inductive ground loops, leading to ringing or even oscillations. The use of an unbroken ground plane on one side of the circuit board can minimize the likelihood of inductive loops within the circuit. However, any particularly sensitive ground-connected nodes in the circuit should be grounded as physically close together as is possible.
2. *Bypass capacitors.* Power supply lines often have significant parasitic inductance and resistance. Large transient load currents can therefore result in voltage spikes on the power supply lines, which can couple onto the signal path within the device. Bypass capacitors are therefore used to lower the impedance of the power supply lines at the point of load, and thus short out the effect of the supply line parasitics. The type of bypass capacitor to use is determined by the application and frequency range of interest. High-speed op-amps work best when their power supply pins are decoupled with RF-quality capacitors.

   Manufacturers often recommend using a composite large-small parallel bypass capacitor with something like a 4.7 uF tantalum capacitor on all supply pins, with a parallel 100 nF ceramic to ensure good capacitive integrity at higher frequencies, where the tantalum becomes inductive. However, a note of caution here: This large-small double capacitor technique relies on the large capacitor having sufficiently high ESR so that at resonance the two capacitors do not create a high-Q parallel filter. In surface-mount designs, a single bypass capacitor may well be better than two due to the inherent high-Q of surface-mount capacitors.

   All bypass capacitor connections should be minimized, since track lengths will simply add more series inductance and resistance to the bypass path. The capacitor should be positioned right next to the power supply pin, with the other lead connected directly to the ground plane.
3. *Sensitive nodes.* Certain nodes within a high-frequency circuit are often sensitive to parasitic components. A current-feedback op-amp, for example, is particularly sensitive to parasitic capacitance at the inverting input, since any capacitance at this point combines with the effective resistance at that node to form a second nondominant pole in the feedback loop. The net result of this additional pole is a reduced phase margin, leading to peaking and even instability. Clearly, great care must be taken during layout to reduce track lengths, etc., at this node. In addition, the stray capacitance to ground at $V(-)$ can be reduced by putting a void area in the ground plane at this point. If the op-amp is used as an inverting amplifier, then the potential of the inverting input is held at virtual ground, and any parasitic capacitance will have less effect. Consequently, the current-feedback op-amp is more stable when used in the inverting rather than the noninverting configuration.
4. *Unwanted oscillations.* Following the preceding guidelines should ensure that your circuit is well behaved. If oscillations still occur, a likely source is unintentional positive feedback due to poor

layout. Output signal paths and other tracks should be kept well away from the amplifier inputs to minimize signal coupling back into the amplifier. Input track lengths should also be kept as short as possible for this same reason.

3.1.9 Broadband Amplifier Stability

Operational amplifiers are generally designed with additional on-chip frequency compensation capacitance in place. This is done to present the applications engineer with an op-amp that is simple to use in negative feedback, with minimal chance of unstable operation. In theory, all will be well, but for three main reasons, op-amps become unstable in the real world of analog electronic circuit design. This section outlines the three main causes for unstable operation of broadband amplifiers and shows practical ways of avoiding these pitfalls.

3.1.9.1 Op-Amp Internal Compensation Strategy

Before dealing with specific stability problems in broadband amplifiers and how to solve them, we will look briefly at the internal frequency compensation strategy used in op-amp design. Generally, op-amps can be classified into two groups, those with two high-voltage gain stages and those with only one stage. The two-stage design provides high open-loop gain but relatively low bandwidth, while the higher speed signal-stage amplifier provides lower open-loop gain but much higher usable bandwidth. Insight into the internal op-amp architecture and the type of compensation used will give the designer valuable information on how to tame the unstable op-amp.

3.1.9.2 Review of the Classical Feedback System

Analyzing the classical feedback system in Figure 3.50 gives the well-known expression for the closed-loop gain, A_c:

$$A_c = A/[1 + B \cdot A] \tag{3.53}$$

where A is the open-loop gain of the amplifier and B the feedback fraction. $T = B \times A$ is referred to as the loop-gain, and the behavior of T over frequency is a key parameter in feedback system design. Clearly, if $T \gg 1$ or $A \gg A_c$, then the closed-loop gain is virtually independent of the open-loop gain A, thus

$$A_c \approx B^{-1} \tag{3.54}$$

This is the most important and desirable feature of negative feedback systems. However, the system will not necessarily be stable as, at higher frequencies, phase lag in the open-loop gain A may cause the feedback to become positive.

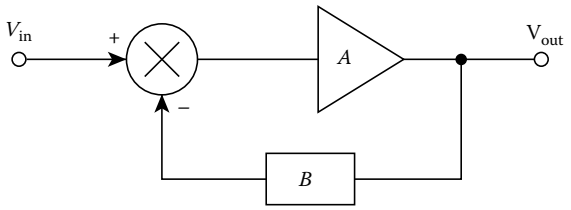


FIGURE 3.50 Classical feedback system.

3.1.9.3 Stability Criteria

Though negative feedback is desirable, it results in potential instability when the feedback becomes positive. The loop-gain T is the best parameter to test whether an amplifier is potentially unstable. The phase margin Φ_M is a common feature of merit used to indicate how far the amplifier is from becoming an oscillator:

$$\Phi_M = 180^\circ + \Phi(|BA| = 1) \tag{3.55}$$

When $\Phi_M = 0°$, the phase of the loop gain, $T = B \times A$ is exactly $-180°$ for $|B \times A| = 1$. The closed-loop gain A_c will become infinite and we have got an oscillator! Clearly, what is required is that $\Phi_M > 0$ and generally the target is to make $\Phi_M \geq 45°$ for reasonably stable performance. However, excessive Φ_M is undesirable if settling time is an important parameter in a particular application.

An op-amp is a general purpose part and so the IC designer strives to produce a maximally versatile amplifier by ensuring that even with 100% feedback, the amplifier circuit will not become unstable. This is done by maintaining a $\Phi_M > 0$ for 100% feedback, that is, when $B = 1$. If the feedback network B is taken to be purely resistive, then any additional phase lag in the loop gain must come from the open-loop amplifier A. Tailoring the phase response of A so that the phase lag is less than 180° up to the point at which $|A| < 1$ or 0 dB ensures that the amplifier is "unconditionally stable"; that is, with any amount of resistive feedback, stable operation is "guaranteed."

Most open-loop op-amps, whether single-stage or two-stage, will exhibit a two-pole response. The separation of these two poles whether at low frequency or high frequency will have a major effect on the stability of the system and it is the op-amp designer's objective to locate these open-loop poles to best advantage to achieve maximum bandwidth, consistent with versatile and stable performance.

3.1.9.4 Two-Stage Op-Amp Architecture

A schematic of the standard two-stage op-amp topology is shown in Figure 3.51. The input differential pair T_1/T_2 provides high gain, as does the second gain stage of T_3/T_4 Darlington pair CE. A high-voltage gain is achieved with this structure, so that output stage is usually a unity voltage gain common-collector output buffer to provide a useful load current drive capability.

The amplifier structure in Figure 3.51 has two internal high-impedance nodes, node X and node Y. These high-impedance nodes are responsible for introducing two dominant poles into the frequency response and their relative location is critical in determining the stability of the amplifier. Each pole contributes a low-pass filter function to the open-loop gain expression of the form

$$[1 + jf/f_P]^{-1} \tag{3.56}$$

Each pole introduces 45° of phase lag at the pole frequency f_P and an additional 45° at $f \approx 10 \times f_P$. With a two-pole amplifier, the open-loop gain A is given by

$$A = A_0/[1 + jf/f_{P1}][1 + jf/f_{P2}] \tag{3.57}$$

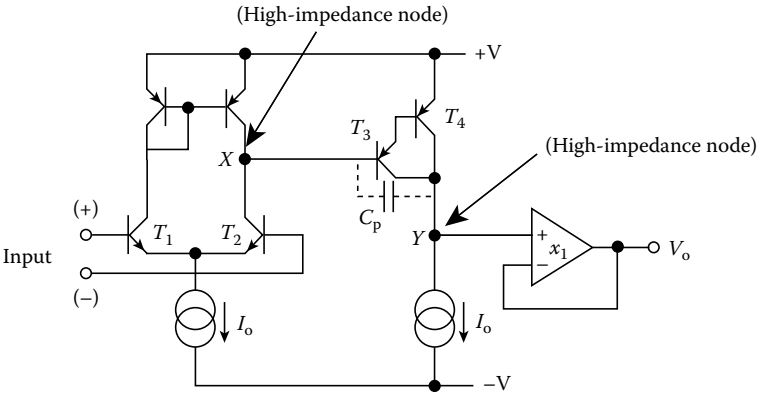


FIGURE 3.51 Architecture of the standard two-stage op-amp.

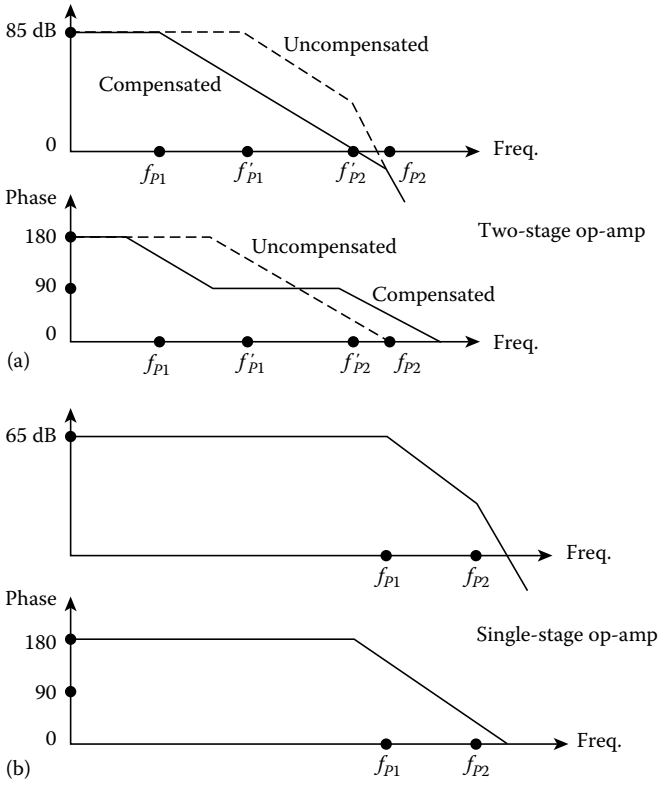


FIGURE 3.52 Pole frequency and phase response for (a) two-stage op-amp and (b) single-stage op-amp.

where A_0 is the dc open-loop gain and f_{P1} and f_{P2} are the two-pole frequencies. A typical plot of A versus f is shown in Figure 3.52a. At low frequencies, where $f \ll f_{P1}$ the gain is flat, and at f_{P1} the gain begins to fall at a rate increasing to −20 dB/decade. The roll-off steepens again at f_{P2} to a final gradient of −40 dB/decade.

It is generally the case that $f_{P1} \ll f_{P2}$ as shown in Figure 3.52a. Turning our attention to the phase plot in Figure 3.52a, at $f = f_{P1}$ the output lags the input by 45°, and as the frequency rises toward f_{P2} the phase lag increases through 135° at f_{P2} to 180° at $f \approx 10 \times f_{P2}$. To ensure unconditionally stable performance, the second pole must be sufficiently far from the first so that the phase margin is large enough.

Figure 3.53 shows curves of the dc value of open-loop gain A_0 versus the ratio N of the pole frequencies ($N = f_{P2}/f_{P1}$) for different values of phase margin. For a given value of $A_0 = 1000$ or +60 dB, the ratio of the pole frequencies must be $N \approx 700$ to obtain a phase margin of 45°.

3.1.9.5 Miller Compensation and Pole Separation

Without any added compensation capacitance, the two open-loop poles of the op-amp are invariably too close to make the amplifier unconditionally stable. The most common compensation method is to add a capacitor between the base and collector of the Darlington pair, shown as C_p in Figure 3.51. This is known as Miller compensation because this strategy makes use of the Miller capacitance multiplication effect discussed earlier. The net result is that the two poles now become significantly far apart, with f_{P1} reducing and f_{P2} increasing, and so the phase margin can be increased to make the op-amp unconditionally stable. However, the penalty of this method is poorer bandwidth and also lower slew rate because of the large capacitance needed, which in practice may be 20 pF or more.

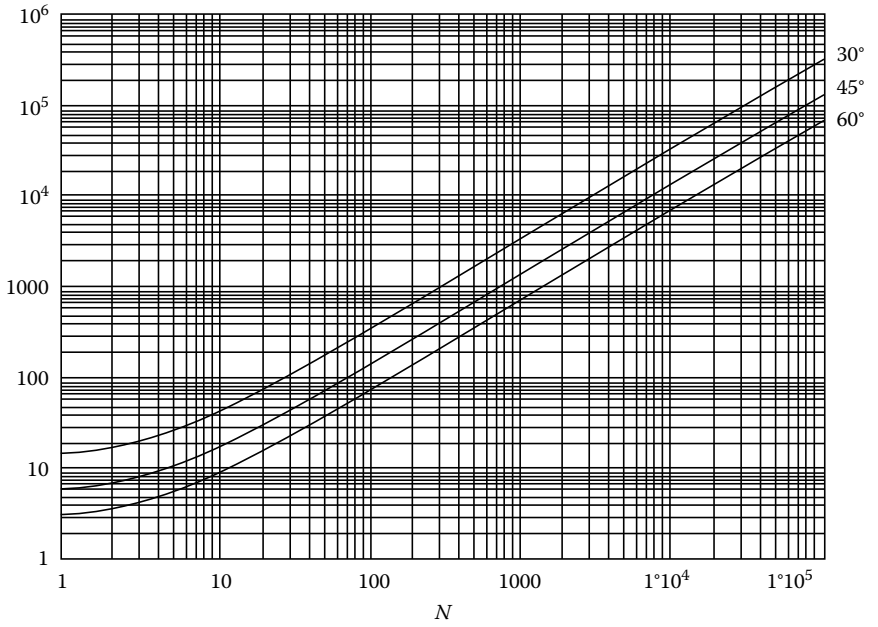


FIGURE 3.53 Low-frequency gain A_0 versus $N\ (=f_{P2}/f_{P1})$ for a two-pole amplifier.

3.1.9.6 Single-Stage Op-Amp Compensation

Figure 3.54 shows a typical simplified circuit schematic of a single-stage op-amp. The input is a differential emitter-coupled pair followed by a folded cascode transistor and an output complementary common-collector buffer. The key difference between this architecture and the two-stage design shown earlier is that X is a low-impedance node, and so the only high-impedance node in the circuit is node Y. Interestingly, the higher frequency nondominant pole of the two-stage amplifier has now become the dominant frequency pole of the single-stage design, as indicated by the second set of curves in Figure 3.52b, which leads to several advantages:

1. The frequency performance off the amplifier is extended. This frequency extension does not lead to a deterioration in phase margin, but simply means that the phase margin problem is shifted up in the frequency domain.

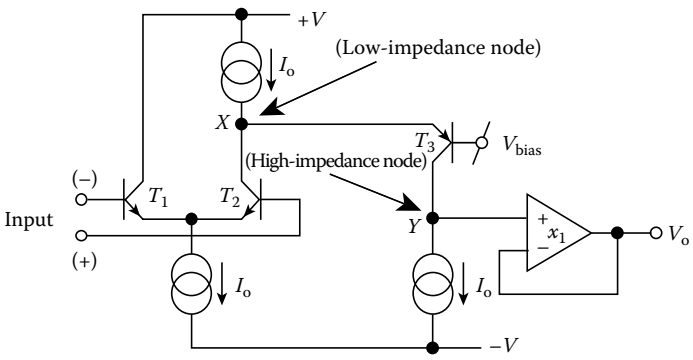


FIGURE 3.54 Architecture of single-stage op-amp.

2. Capacitance at the high-impedance Y node reduces bandwidth, but now improves phase margin.
3. A single value of a few pFs of grounded capacitor at Y will now act as a satisfactory compensation capacitor, unlike the large Miller capacitor required in the two-stage design.
4. The slewing capability of this single-stage structure is very good as a result of the much smaller compensation capacitor.
5. Clearly, it is much more straightforward to develop a stable amplifier for high-frequency applications if it has essentially only one voltage gain stage and so high-frequency op-amp designers generally opt for a single gain stage architecture.

3.1.9.7 Grounded Capacitor Compensation

Typical A_{OL} versus f responses of two single-stage op-amps are shown in Figure 3.55, indicating one high-frequency pole and its proximity to the nondominant pole.

The curves are taken from data for (a) a 2 GHz gain-bandwidth product voltage-feedback op-amp and (b) a 150 MHz current-feedback op-amp. In both cases, the phase characteristics demonstrate the expected 45° lag at the pole frequency, and the slow roll-off in phase at high frequency due to the presence of the very-high-frequency poles.

Both single-stage and two-stage op-amps can be approximated by the two-pole macromodel shown in Figure 3.56. Transconductance G_M and output resistance R_0 represent the gain per stage of $G_M \times R_0$. The difference between the two-stage and single-stage op-amp models is that R_{01} of the single-stage is of the order of $[G_M]^{-1}$ and the dominant compensation capacitor is C_2. C_P in the case of the single stage will

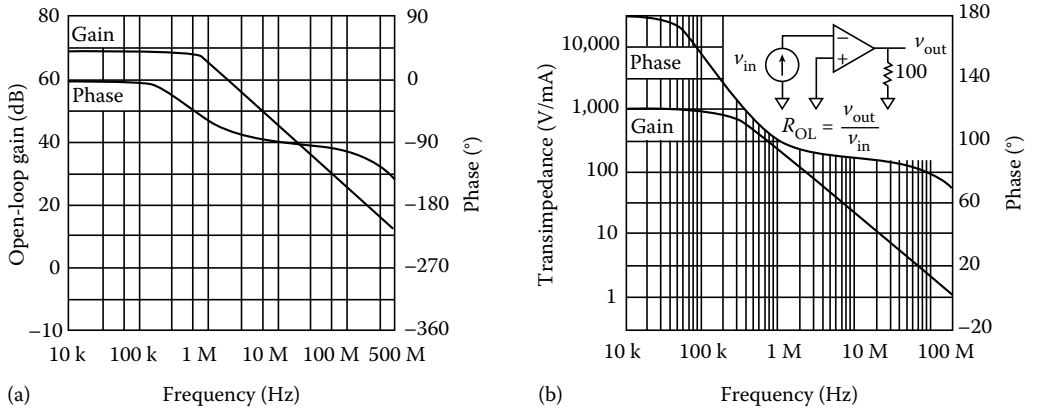


FIGURE 3.55 Single-pole op-amps; open-loop gain and phase frequency characteristics. (a) Voltage feedback. (b) Current feedback.

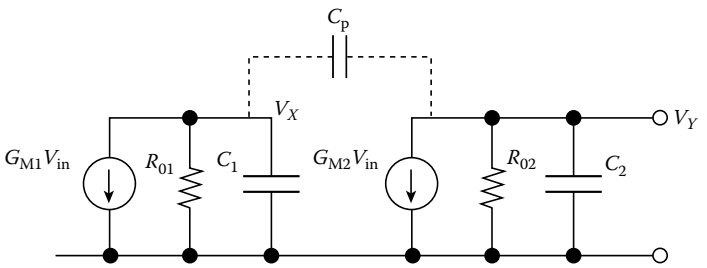


FIGURE 3.56 Partial equivalent circuit of two-pole op-amp.

simply be a feedback parasitic capacitor, while in the case of a two-stage it will be the dominant Miller compensating capacitor. This simple model is an excellent first-cut tool for determining pole locations, and the value of compensation capacitor for a desired bandwidth and stability.

3.1.9.8 High-Frequency Performance

Although the bandwidth in a single-stage design is significantly extended, circuit parasitics become more important. We are confronted with the problem of potential instability, since at higher frequencies the "working environment" of the op-amp becomes very parasitic sensitive; in other words, now op-amp-embedded parasitics cannot be neglected.

An op-amp in closed-loop can be considered at three levels, as shown schematically in Figure 3.57. The inner triangle is the ideal op-amp, internally compensated by the op-amp designer for stable operation using the circuit techniques outlined earlier. High-frequency amplifiers are sensitive to parasitics of the surrounding circuit. The key parasitics within the outer triangle include power supply lead inductance, stray capacitance between power supply pins, and input to ground capacitance. The effect of these parasitics is to destabilize the amplifier, and so the designer is confronted with the task of reestablishing stable operation. The approach needed to achieve this parallels the work of the op-amp designer. The parasitics almost always introduce additional extrinsic nondominant poles, which need to be compensated. The task of compensation cannot be attempte without considering the outer or third level, which includes the closed-loop gain defining components together with the load impedance. Again, stray reactance associated with these components will modify the loop gain, and so to guarantee stable operation of the closed-loop amplifier it is necessary to compensate the complete circuit.

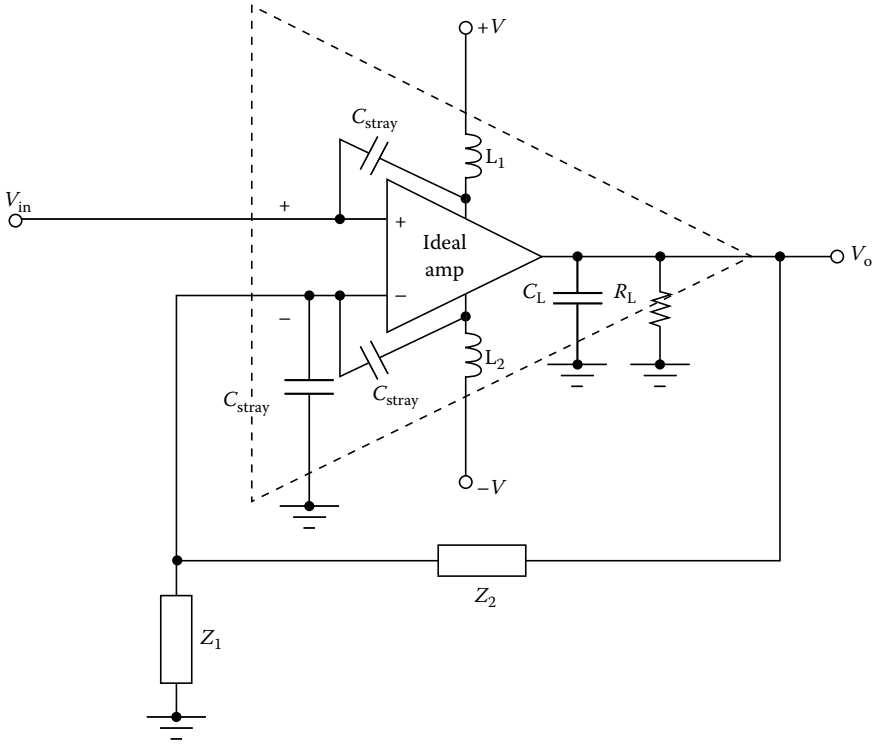


FIGURE 3.57 Real feedback amplifier.

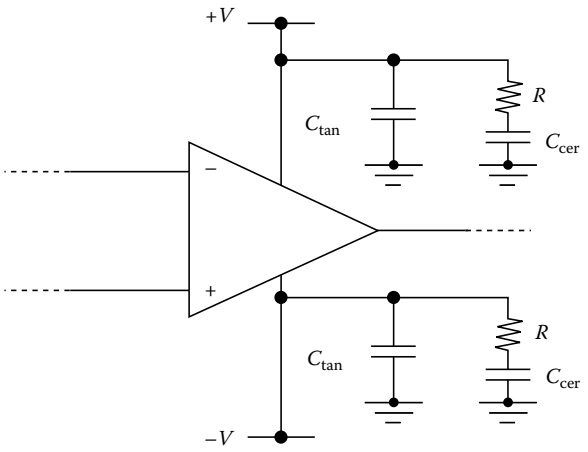


FIGURE 3.58 Supply decoupling circuitry (C_{CER} = ceramic capacitor and C_{TAN} = tantalum).

3.1.9.9 Power Supply Impedance

In this section, we consider the ways in which the impedance of the power supply can affect the frequency response of the amplifier. First, some important rules are

1. There is no such thing as an ideal zero-impedance power supply.
2. Real power supplies have series *R–L* impedance and at high frequencies the inductance matters most.
3. Power supply inductance causes "bounce" on the power supply voltage, generating unwanted feedback via parasitic capacitive links to the inputs. Power supply "bounce" increases with increasing load current.
4. Supply decoupling capacitors act as "short-term local batteries" to maintain power supply integrity, and it is important that they are placed as close as possible to the power supply pins of the op-amp.

Large electrolytic capacitors are fine at low frequencies but are inductive at high frequencies. Figure 3.58 shows commonly used decoupling circuitry. Small-sized tantalum electrolytics are preferred, while a parallel ceramic capacitor with low series inductance takes over the decoupling role at high frequencies. The added series *R* prevents the inductance of the electrolytic resonating with the ceramic capacitor. The waveforms in Figure 3.59 illustrate the benefits of good decoupling.

3.1.9.10 Effects of Resistive and Capacitive Loads

The load presented to an amplifier is likely to have both resistive and capacitive components, as illustrated previously in Figure 3.57. Increasing the load current causes power supply ripple, so good power supply decoupling is vital.

A closed-loop amplifier with voltage-sampled negative feedback results in a very-low output impedance, so it is natural to think that the effects of any load would be shunted out by this low impedance. In reality, the load has an important effect on the amplifier and must not be overlooked. Resistive loads, for example, cause two main effects. First, as a voltage divider with the open-loop output resistance of the op-amp r_0, the open-loop gain is reduced. This effect is small unless the load resistance approaches r_0. Second, the load current is routed to the output pin via the supply pins, and as the load current increases,

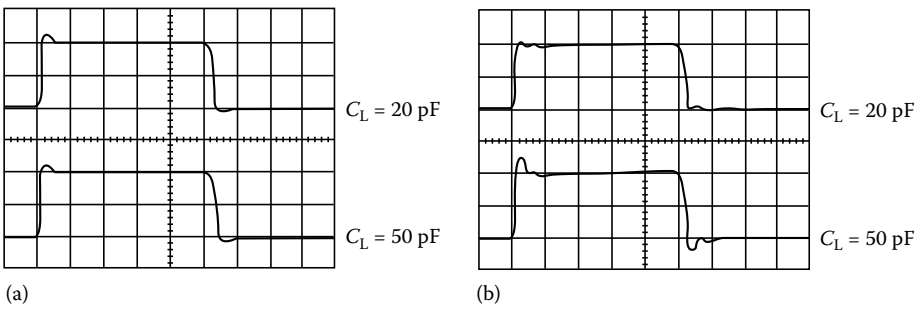


FIGURE 3.59 High-speed voltage buffer: (a) with and (b) without supply decoupling.

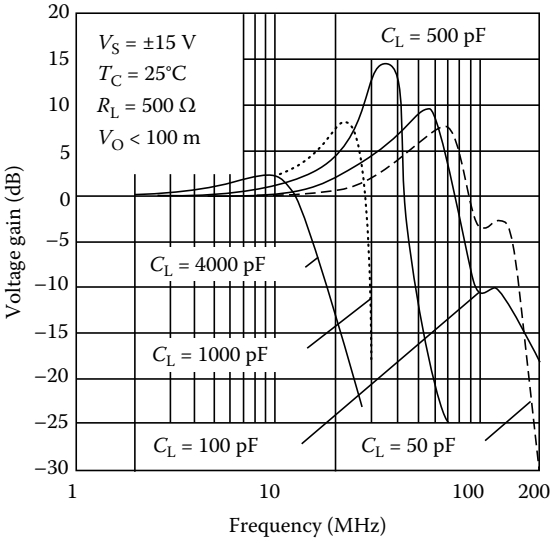


FIGURE 3.60 Load capacitance causes gain peaking.

the supply pin voltage is modulated. This effect is more important, since the integrity of the power supply will be degraded. Again, good supply decoupling is essential to minimize this effect.

Capacitive load current is proportional to the derivative of output voltage, and the maximum capacitive output current demand occurs when dV_{out}/dt is a maximum. Though not directly a stability issue, the designer must remember that a capacitive load demands high-output current at high frequencies and at high amplitude, that is,

$$I_{max} = C_L \cdot 2\pi f_{max} \cdot V_{outpeak} \tag{3.58}$$

Figure 3.60 illustrates the effect of load capacitance on the loop gain.

C_1 together with the equivalent output resistance of the op-amp adds an additional pole into the loop gain of the form

$$V_F/V_{out} = B = 1/[1 + jf/f_L] \quad \text{where } f_L = 1/2\pi r_0 \cdot C_L \tag{3.59}$$

The load resistance has a minor influence on the loop gain compared to the effects of load capacitance by slightly reducing the value of dc open-loop gain by factor K, where $K = R_L/[r_0 + R_L]$, as described above. Since the effective output resistance reduces to $r_{0'} = r_0/R_L$, then f_L changes to $f_{L'} = 1/2\pi r_{0'} C_L$.

3.1.9.11 Neutralizing the Phase Lag

To compensate for high-frequency phase lag, the simplest technique is to add a series resistance R between the output of the op-amp and the load connection point, as shown in Figure 3.61.

The series resistor adds a zero into the V_F/V_{out} equation, which changes to

$$V_F/V_{out} = K \cdot [1 + jf/f_Z]/[1 + jf/f_P] \tag{3.60}$$

where $K = [R + R_L]/[r_0 + R + R_L]$, $f_P = 1/[2\pi(r_0 + R)/R_L \cdot C_L]$ and $f_Z = 1/[2\pi R_{L//R} \cdot C_L] = f_P \cdot [1 + r_0/R]$, so clearly; $f_P < f_Z$.

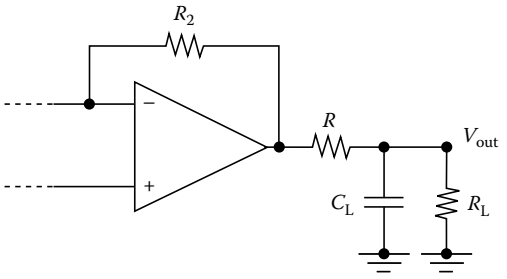


FIGURE 3.61 Load capacitance neutralization.

The phase lag introduced by the pole is compensated by the phase lead of the zero at higher frequencies. The maximum phase lag is limited if the zero is close to the pole, almost eliminating the effects of the load capacitor. Maximum phase lag in V_F/V_{out} occurs at $f = f_M$, where f_M is given by

$$f_M = [f_P \cdot f_Z]^{1/2} = f_P \times (1 + r_0/R)^{1/2} \tag{3.61}$$

and at f_M the phase lag $\Phi = \Phi'$ is given by

$$\begin{aligned} \Phi' &= 90^\circ - 2 \cdot \tan^{-1}[f_M/f_P] = 90^\circ - 2 \cdot \tan^{-1}\left[(1 + r_0/R)^{1/2}\right] \\ \Phi' &\approx -19.5^\circ \quad \text{for} \quad R = r_0 \\ \Phi' &\approx -8.2^\circ \quad \text{for} \quad R = 2 \cdot r_0 \\ \Phi' &\approx -6.4^\circ \quad \text{for} \quad R = 3 \cdot r_0 \end{aligned} \tag{3.62}$$

These values show that the added lag Φ' is not excessive as long as $R > r_0$. The disadvantage with this method is that the series resistor is in direct line with the output current, increasing the output resistance of the amplifier and limiting the output current drive capability. The output impedance also goes inductive at high frequencies.

An alternative way of solving the problem of capacitive load is to view the closed-loop output resistance of the op-amp as being inductive, since the closed-loop output impedance of the op-amp is essentially the open-loop output resistance divided by the loop gain. As the loop gain falls with frequency, the output impedance rises, and thus appears inductive. Adding a load capacitor generates a resonant circuit. The solution is to "spoil" the Q of the resonator, therefore minimizing the added phase lag of C_L.

Adding a so-called series R–C "snubber," as in Figure 3.62, effects a cure. The resistor R is ac coupled by the capacitor at high frequencies and spoils the Q. Effectively, C_L resonates with the inductive output impedance, and at this frequency leaves the R–C snubber as a "new" load. The equivalent circuit is therefore close to the previous compensation method shown in Figure 3.61, but with the added advantage that now the load current is not carried by the series resistance. To select the snubber component values, make $R = 1/2\pi f_0 C$, where f_0 is the resonant frequency, which can simply be determined experimentally from the amplifier without the snubber in place. The value of the series capacitance is a compromise: too big and it will increase the effective load capacitance. Choosing $C = C_L$ works reasonably well in practice.

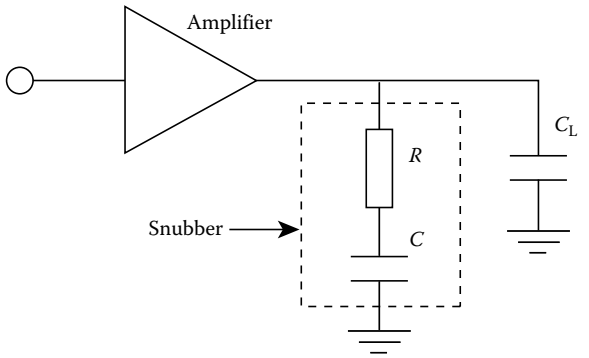


FIGURE 3.62 Snubber cures capacitive load peaking.

3.1.9.12 Inverting Input Capacitance to Ground

With most broadband bipolar op-amps, parasitic capacitance to ground adds an additional pole (and hence phase lag) into the feedback path, which threatens stability. Stray capacitance C_1 at the inverting input pin (shown previously in Figure 3.57) modifies B and adds phase lag in the loop-gain T, compromising stability.

Solving for B with C_1 taken into account will clarify the problem. It is simple to show that

$$B = V_F/V_{out} = Z_1/[Z_1 + Z_2] \tag{3.63}$$

where $Z_1 = R_1/[1 + j\omega R_1 C_1]$ and $Z_2 = R_2$. Substituting, we get

$$B = K/[1 + jf/f_C] \tag{3.64}$$

where $K = R_1\ [R_1 + R_2]$ and $f_C = 1/[2\pi C_1 R_1/R_2]$.

The additional pole at $f = f_C$ will now give the circuit a very undesirable three-pole loop gain, which could cause significant gain peaking, as shown in Figure 3.63. f_C could be made high by choosing relatively low values of $R_1 // R_2$ but the additional pole can be eliminated by adding a feedback capacitor C_2 across resistor R_2 to give pole-zero cancellation.

$$Z_1 = R_1/[1 + j\omega R_1 C_1] \quad \text{and} \quad Z_2 = R_2/[1 + j\omega R_2 C_2] \tag{3.65}$$

If $R_1C_1 = R_2C_2$, then $B = Z_1/[Z_1 + Z_2] = R_1/[R_1 + R_2]$, making B frequency independent. The design equation for C_2 is then

$$C_2 = C_1 \cdot R_1/R_2 \tag{3.66}$$

If the open-loop phase margin Φ_M needs to be increased for the desired value of closed-loop gain, and the inverting capacitance C_1 has its inevitable high-frequency influence, then the optimum solution for C_2 would be to locate the zero on the second pole of the loop-gain response following the procedure given above.

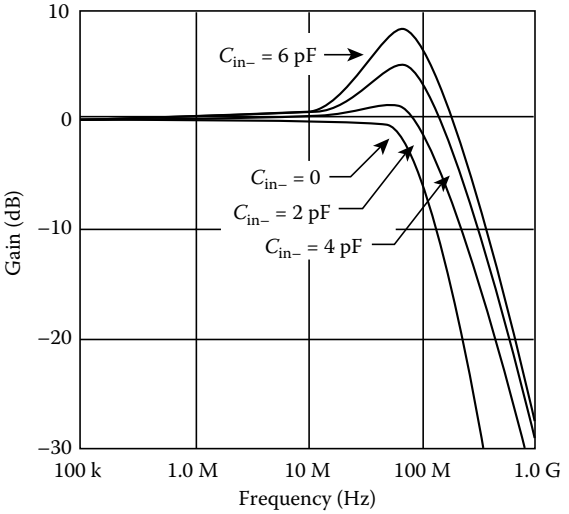
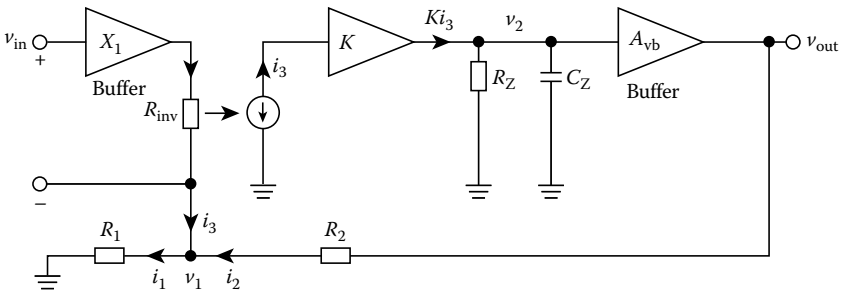


FIGURE 3.63 Stray input capacitance causes gain peaking.

3.1.10 Conclusions

This chapter hopefully serves to illustrate some of the modern techniques the practicing engineer will encounter when designing broadband bipolar amplifiers. It focuses mainly upon key generic building blocks and methodologies for broadband design. Many circuits and design techniques have not been covered, but analysis techniques described should serve as a foundation for the analysis of other broadband designs. Furthermore, comprehensive analytical treatment of many alternative broadband bipolar circuits can be found in the texts [6,13–15].

Appendix A: Transfer Function and Bandwidth Characteristic of Current-Feedback Operational Amplifier



$$-i_1 + i_2 + i_3 = 0 \tag{3.67}$$

$$i_1 = \frac{v_1}{R_1} \tag{3.68}$$

$$i_2 = \frac{v_{out} - v_1}{R_2} \tag{3.69}$$

$$i_3 = \frac{v_{in} - v_1}{R_{inv}} \tag{3.70}$$

$$v_2 = \frac{Ki_3R_Z}{1 + j\omega R_Z C_Z} \tag{3.71}$$

$$v_{out} = A_{vb}v_2 \tag{3.72}$$

Substituting Equations 3.68 through 3.70 into Equation 3.67 yields

$$-\frac{v_1}{R_1} + \frac{v_{out} - v_1}{R_2} + \frac{v_{in} - v_1}{R_{inv}} = 0$$

Rearranging for v_1 gives

$$v_1 = \frac{v_{in}R_2/R_{inv} + v_{out}}{1 + (R_2/R_1) + (R_2/R_{inv})}$$

From Equations 3.71 and 3.72, it is clearly seen that

$$v_{out} = \frac{A_{vb} K i_3 R_Z}{1 + j\omega R_Z C_Z} \tag{3.73}$$

Substituting for i_1 and i_2 from Equations 3.68 and 3.69 into Equation 3.67 gives

$$i_3 = v_1 \left(\frac{1}{R_1} + \frac{1}{R_2}\right) - \frac{v_{out}}{R_2}$$

Substitute for v_1:

$$i_3 = \left[\frac{v_{in} R_2/R_{inv} + v_{out}}{1 + (R_2/R_1) + (R_2/R_{inv})}\right]\left(\frac{1}{R_1} + \frac{1}{R_2}\right) - \frac{v_{out}}{R_2}$$

Substitute for i_3 from Equation 3.73:

$$\frac{v_{out}(1 + j\omega R_Z C_Z)}{A_{vb} K R_Z} = \left\{\left[\frac{v_{in} R_2/R_{inv} + v_{out}}{1 + (R_2/R_1) + (R_2/R_{inv})}\right]\left(\frac{1}{R_1} + \frac{1}{R_2}\right) - \frac{v_{out}}{R_2}\right\}$$

rearranging

$$v_{out}\left[\frac{(1 + j\omega R_Z C_Z)}{A_{vb} K R_Z} - \frac{(1/R_1 + 1/R_2)}{1 + (R_2/R_1) + (R_2/R_{inv})} + \frac{1}{R_2}\right] = \frac{(v_{in} R_2/R_{inv})(1/R_1) + (1/R_2)}{1 + (R_2/R_1) + (R_2/R_{inv})}$$

$$\frac{v_{out}}{v_{in}} = \frac{1 + (R_2/R_1)}{\frac{R_{inv}(1 + (R_2/R_1) + (R_2/R_{inv}))(1 + j\omega R_Z C_Z)}{A_{vb} K R_Z} - R_{inv}((1/R_1) + (1/R_2)) + \frac{R_{inv}(1 + (R_2/R_1) + (R_2/R_{inv}))}{R_2}}$$

$$\frac{v_{out}}{v_{in}} = \frac{1 + (R_2/R_1)}{\frac{R_{inv}(1 + (R_2/R_1)) + R_2}{A_{vb} K R_Z} + \frac{(R_{inv}(1 + (R_2/R_1)) + R_2) j\omega R_Z C_Z}{A_{vb} K R_Z} + 1}$$

Factorize the denominator

$$\frac{v_{out}}{v_{in}} = \frac{1 + (R_2/R_1)}{\left[1 + \frac{R_{inv}(1 + (R_2/R_1)) + R_2}{A_{vb} K R_Z}\right]\left[1 + \frac{\frac{(R_{inv}(1 + (R_2/R_1)) + R_2) j\omega R_Z C_Z}{A_{vb} K R_Z}}{1 + \frac{R_{inv}(1 + (R_2/R_1)) + R_2}{A_{vb} K R_Z}}\right]}$$

$$\frac{v_{out}}{v_{in}} = \frac{1 + (R_2/R_1)}{\left[1 + \frac{R_{inv}(1 + (R_2/R_1)) + R_2}{A_{vb} K R_Z}\right]\left[1 + j\omega C_Z\left\{\frac{R_{inv}(1 + (R_2/R_1)) + R_2}{A_{vb} K + \frac{R_{inv}(1 + (R_2/R_1)) + R_2}{R_Z}}\right\}\right]}$$

If we assume that R_Z is very large, then

$$\frac{R_{inv}(1 + (R_2/R_1)) + R_2}{R_Z} \approx 0$$

and the transfer function becomes

$$\frac{v_{out}}{v_{in}} = \frac{1 + (R_2/R_1)}{1 + j\omega C_Z\left[\frac{R_{inv}(1 + (R_2/R_1)) + R_2}{A_{vb} K}\right]}$$

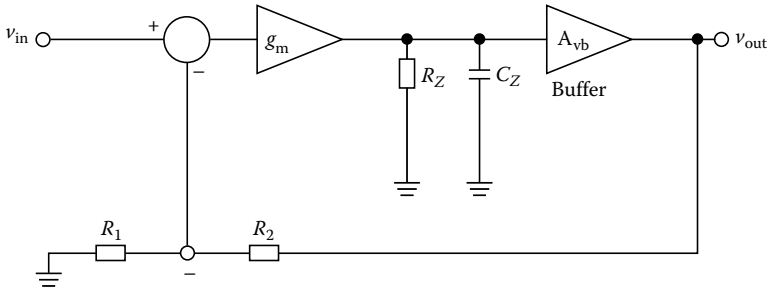
The pole frequency is given by

$$f_{-3\text{ dB}} = \frac{A_{\text{vb}}K}{2\pi C_Z[R_{\text{inv}}(1 + (R_2/R_1)) + R_2]}$$

The gain-bandwidth product is given by

$$GBW = \frac{A_{\text{vb}}K[1 + (R_2/R_1)]}{2\pi C_Z[R_{\text{inv}}(1 + (R_2/R_1)) + R_2]}$$

Appendix B: Transfer Function and Bandwidth Characteristic of Voltage-Feedback Operational Amplifier



$$\begin{aligned}
v_{\text{out}} &= \left(v_{\text{in}} - \frac{R_1}{R + R_2 v_{\text{out}}}\right)\frac{g_{\text{m}}R_Z A_{\text{vb}}}{1 + j\omega R_Z C_Z} \\
v_{\text{out}}&\left[1 + \frac{R_1 g_{\text{m}} A_{\text{vb}} R_Z}{(R_1 + R_2)(1 + j\omega R_Z C_Z)}\right] = v_{\text{in}}\frac{g_{\text{m}}A_{\text{vb}}R_Z}{1 + j\omega R_Z C_Z} \\
\frac{v_{\text{out}}}{v_{\text{in}}} &= \frac{g_{\text{m}}A_{\text{vb}}R_Z/(1 + j\omega R_Z C_Z)}{1 + \frac{R_1 g_{\text{m}} A_{\text{vb}} R_Z}{(R_1 + R_2)(1 + j\omega R_Z C_Z)}}
\end{aligned} \tag{3.74}$$

Multiply the numerator and denominator by $(1 + j\omega R_Z C_Z)/g_{\text{m}}\, A_{\text{vb}}\, R_Z$

$$\begin{aligned}
\frac{v_{\text{out}}}{v_{\text{in}}} &= \frac{1}{\frac{1 + j\omega R_Z C_Z}{g_{\text{m}}A_{\text{vb}}R_Z} + \frac{R_1}{(R_1 + R_2)}} \\
\frac{v_{\text{out}}}{v_{\text{in}}} &= \frac{1 + (R_2/R_1)}{\left[\frac{1 + j\omega R_Z C_Z}{g_{\text{m}}A_{\text{vb}}R_Z}\right][1 + (R_2/R_1)] + 1} \\
\frac{v_{\text{out}}}{v_{\text{in}}} &= \frac{1 + (R_2/R_1)}{\frac{1 + (R_2/R_1)}{g_{\text{m}}A_{\text{vb}}R_Z} + \frac{j\omega R_Z C_Z(1 + (R_2/R_1))}{g_{\text{m}}A_{\text{vb}}R_Z} + 1}
\end{aligned}$$

get $1 + [1 + (R_2/R_1)/g_{\text{m}}\, A_{\text{vb}}\, R_Z]$ out of the denominator

$$\frac{v_{\text{out}}}{v_{\text{in}}} = \frac{1 + (R_2/R_1)}{\left[1 + \frac{1 + (R_2/R_1)}{g_m A_{vb} R_Z}\right]\left[1 + \frac{\frac{j\omega R_Z C_Z (1 + (R_2/R_1))}{g_m A_{vb} R_Z}}{1 + \frac{1 + (R_2/R_1)}{g_m A_{vb} R_Z}}\right]}$$

multiply the denominator bracket by $g_m\, A_{vb}\, R_z/[1 + (R_2/R_1)]$

$$\frac{v_{\text{out}}}{v_{\text{in}}} = \frac{1 + (R_2/R_1)}{\left[1 + \frac{1 + (R_2/R_1)}{g_m A_{vb} R_Z}\right]\left[1 + \frac{j\omega R_Z C_Z (1 + (R_2/R_1))}{g_m A_{vb} R_Z + (1 + (R_2/R_1))}\right]}$$

$$\frac{v_{\text{out}}}{v_{\text{in}}} = \frac{1 + (R_2/R_1)}{\left[1 + \frac{1 + (R_2/R_1)}{g_m A_{vb} R_Z}\right]\left[1 + \frac{j\omega R_Z C_Z}{1 + \frac{g_m A_{vb} R_Z}{1 + (R_2/R_1)}}\right]}$$

assuming that $g_m\, A_{vb} R_Z$ is much larger than $1 + R_2/R_1$, then

$$\frac{v_{\text{out}}}{v_{\text{in}}} = \frac{1 + (R_2/R_1)}{1 + j\omega\left[\frac{R_Z C_Z}{1 + \frac{g_m A_{vb} R_Z}{1 + (R_2/R_1)}}\right]}$$

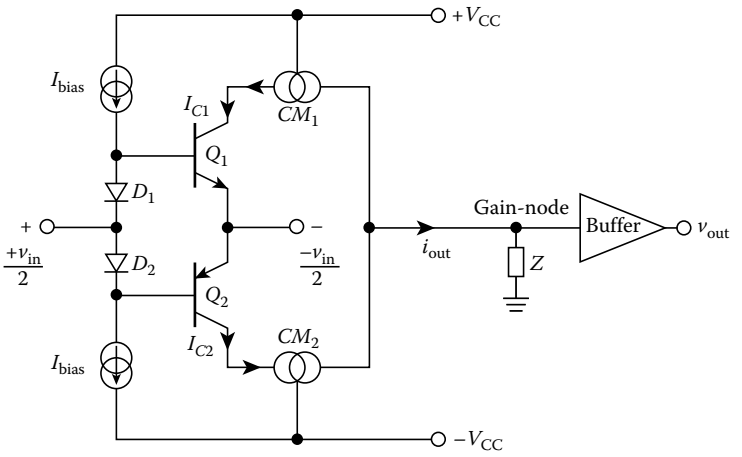
The pole frequency is given by

$$f_{-3\text{ dB}} = \frac{1 + \frac{g_m A_{vb} R_Z}{1 + (R_2/R_1)}}{2\pi R_Z C_Z}$$

The gain-bandwidth product is given by

$$GBW = \frac{(1 + (R_2/R_1))\left[1 + \frac{g_m A_{vb} R_Z}{(1 + (R_2/R_1))}\right]}{2\pi R_Z C_Z}$$

Appendix C: Transconductance of the Current-Feedback Op-Amp Input Stage



$$v_{\text{in}} = v_1 - v_2$$

$$i_{\text{out}} = I_{CI} - I_{C2}$$

$$I_{C1} = I_{S1} e^{\frac{v_{BE1}}{v_T}}$$

$$I_{C2} = I_{S2} e^{\frac{v_{BE2}}{v_T}}$$

$$V_{BE1} = V_{DQ1} + v_{\text{in}}$$

$$V_{BE2} = v_{\text{in}} - V_{DQ2}$$

$$I_{C1} = I_{S1} e^{\left(\frac{V_{DQ1}}{V_T} + \frac{v_{\text{in}}}{v_T}\right)}$$

$$I_{C2} = I_{S2} e^{\left(\frac{V_{DQ1}}{V_T} - \frac{v_{\text{in}}}{v_T}\right)}$$

$$I_{C1} = I_{CQ1} e^{\frac{v_{\text{in}}}{v_T}}$$

$$I_{C2} = I_{CQ2} e^{-\frac{v_{\text{in}}}{v_T}}$$

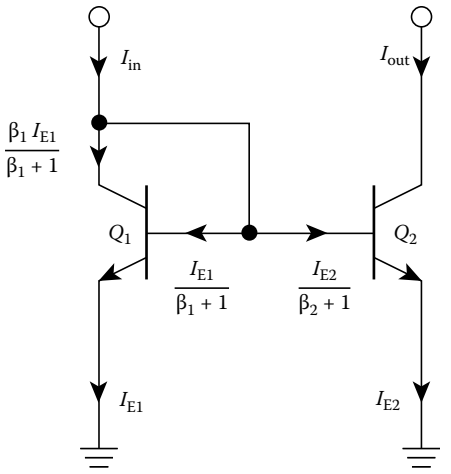
Assuming matched transistors then, $I_{\text{CQ1}} = I_{\text{CQ2}} = I_{\text{CQ}}$

$$i_{\text{out}} = I_{\text{C1}} - I_{\text{C2}} = I_{\text{CQ}} \left[e^{+\left(\frac{v_{\text{in}}}{v_{\text{T}}}\right)} - e^{-\left(\frac{v_{\text{in}}}{v_{\text{T}}}\right)} \right]$$

$$\frac{i_{\text{out}}}{I_{\text{CQ}}} = y = \left[e^{x} - e^{-x}\right] = 2 \sinh(x)$$

where $x = + v_{\text{in}}/V_{\text{T}}$.

Appendix D: Transfer Function of Widlar Current Mirror



$$I_{in} = I_{E1} + \frac{I_{E2}}{\beta_2 + 1}$$

$$I_{in} = \frac{I_{E1}(\beta_2 + 1) + I_{E2}}{\beta_2 + 1}$$

$$I_{out} = \beta_2 I_{B2}$$

$$I_{out} = \frac{\beta_2 I_{E2}}{\beta_2 + 1}$$

$$\frac{I_{out}}{I_{in}} = \frac{(\beta_2 I_{E2})(\beta_2 + 1)}{(\beta_2 + 1)[I_{E1}(\beta_2 + 1) + I_{E2}]}$$

$$\frac{I_{out}}{I_{in}} = \frac{\beta_2 I_{E2}}{I_{E1}(\beta_2 + 1) + I_{E2}}$$

$$\frac{I_{out}}{I_{in}} = \frac{1}{\frac{I_{E1}(\beta_2 + 1)}{I_{E2}\beta_2} + \frac{1}{\beta_2}}$$

For

$$\frac{I_{E1}}{I_{E2}} = \frac{I_{S1}\left(\frac{\beta_1 + 1}{\beta_1}\right) e^{\frac{v_{BE1}}{V_T}}}{I_{S2}\left(\frac{\beta_2 + 1}{\beta_2}\right) e^{\frac{v_{BE2}}{v_T}}}$$

Then, as $V_{BE1} = V_{BE2}$,

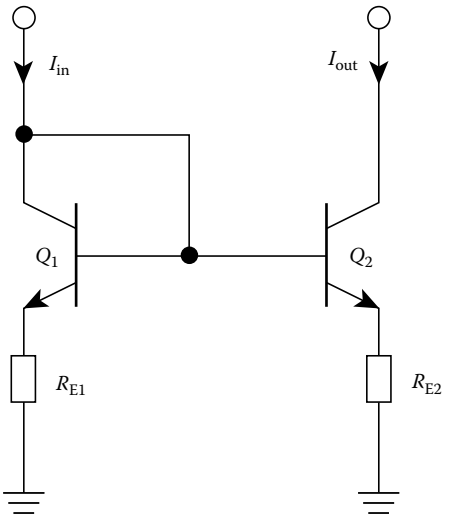
$$\frac{I_{E1}}{I_{E2}} = \frac{I_{S1}(\beta_1 + 1/\beta)}{I_{S2}(\beta_2 + 1/\beta_2)}$$

$$\frac{I_{out}}{I_{in}} = \frac{1}{\frac{I_{S1}(\beta_1 + 1)}{I_{S2}\beta_1} + \frac{1}{\beta_2}}$$

Assume $\beta_1 = \beta_2 = \beta$, $I_{S1} = I_{S2}$. Then

$$\frac{I_{out}}{I_{in}} = \frac{\beta}{\beta + 2}$$

Appendix E: Transfer Function of Widlar Current Mirror with Emitter Degeneration Resistors



Assuming that $\beta \gg 1$, then

$$V_{BE1} + I_{in}R_1 = V_{BE2} + I_{out}R_2$$

$$I_{out} = \frac{I_{in}R_1}{R_2} + \frac{(V_{BE1} - V_{BE2})}{R_2}$$

$$\frac{I_{out}}{I_{in}} = \frac{R_1}{R_2} + \frac{(V_{BE1} - V_{BE2})}{I_{in}R_2}$$

$$\frac{I_{out}}{I_{in}} = \frac{R_1}{R_2} + \frac{V_T \ln\left(\frac{I_{in}}{I_{S1}}\frac{I_{S2}}{I_{out}}\right)}{I_{in}R_2}$$

$$\frac{I_{out}}{I_{in}} = \frac{R_1}{R_2} + \frac{V_T(\ln(I_{in}/I_{out}) + (\Delta V_{BE}/V_T))}{I_{in}R_2}$$

$$\frac{I_{out}}{I_{in}} = \frac{R_1}{R_2} + \frac{V_T(\ln(I_{in}/I_{out}))}{I_{in}R_2} + \frac{\Delta V_{BE}}{I_{in}R_2}$$

Assuming that the term $V_T[\ln(I_{in}/I_{out})]/I_{in}R_2$ is small compared with the other terms, then

$$\frac{I_{out}}{I_{in}} = \frac{R_1}{R_2} + \frac{\Delta V_{BE}}{I_{in}R_2}$$

References

1. K. C. Smith and A. S. Sedra, The current conveyor—A new circuit building block, *Proc. IEEE*, 56, 1368–1369, 1968.
2. A. Sedra and K. C. Smith, A second generation current-conveyor and its applications, *IEEE Trans. Circuit Theory*, CT-17, 132–134, 1970.

3. B. Wilson, High performance current conveyor implementation, *Electron. Lett*., 20(24), 990–991, 1984.
4. C. Toumazou, F. J. Lidgey, and C. Makris, Extending voltage-mode op-amps to current-mode performance, *Proc. IEE: Pt. G*, 137(2), 116–130, 1990.
5. PA630 Data Sheet, Photronics Co., Ottawa, PQ, Canada.
6. C. Toumazou, F. J. Lidgey, and D. Haigh, Eds., *Analogue IC Design—The Current-Mode Approach*, Exeter, England: Peter Peregrinus, 1990.
7. CCII01 Data Sheet, LTP Electronics, Headington, Oxford, England.
8. D. F. Bowers, A precision dual current-feedback operational amplifier, in *Proc. IEEE Bipolar Circuits Technol. Meet*., Minneapolis, MN, Sep. 1988, pp. 68–70.
9. D. F. Bowers, Applying current feedback to voltage amplifier, in *Analogue IC Design: The Current-Mode Approach*, edited by C. Toumazou, F. J. Lidgey, and D. G. Haigh, Eds. Exeter, England: Peter Peregrinus, 1990, Chap. 16, pp. 569–595.
10. I. A. Koullias, A wideband low-offset current-feedback op amp design, in *Proc. IEEE* 1989 *Bipolar Circuits Technol. Meet*., Minneapolis, MN, Sep. 18–19, 1989, pp. 120–123.
11. A. Payne and C. Toumazou, High frequency self-compensation of current feedback devices, in *Proc. IEEE ISCAS*, San Diego, CA, May 10–13, 1992, pp. 1376–1379.
12. T. Vanisri and C. Toumazou, Wideband and high gain current-feedback op-amp, *Electron. Lett*., 28 (18), 1705–1707, 1992.
13. A. Grebene, *Bipolar and MOS Analog Integrated Circuit Design*, New York: Wiley, 1984.
14. C. Toumazou, Ed., *Circuits and Systems Tutorials*, New York: IEEE ISCAS, 1994.
15. *High Performance Analog Integrated Circuits*. Élantec Data Book, Elantec (Intersil Corporation, Milpitas, CA), 1994.

3.2 Bipolar Noise

Alicja Konczakowska and Bogdan M. Wilamowski

Bipolar transistors and other electronic devices generate inherent electrical noise. This limits the device operation at a small-signal range. There are a few different sources of noise, such as thermal noise, shot noise, generation–recombination, $1/f$ (flicker noise), and $1/f^2$ noise, burst noise or random telegraph signal noise (RTS noise), and avalanche noise [1,6].

3.2.1 Thermal Noise

Thermal noise is created by random motion of charge carriers due to the thermal excitation [1]. This noise is sometimes known as the Johnson noise. In 1905 Einstein presented his theory of fluctuating movement of charges in thermal equilibrium. This theory was experimentally verified by Johnson in 1928. The thermal motion of carriers creates a fluctuating voltage on the terminals of each resistive element. The average value of this voltage is zero, but the power on its terminals is not zero. The internal noise voltage source or current source is described by Nyquist equation

$$\overline{v_{\mathrm{n}}^2} = 4kTR\Delta f, \quad \overline{i_{\mathrm{n}}^2} = \frac{4kT\Delta f}{R} \tag{3.75}$$

where
k is the Boltzmann constant
T is absolute temperature
$4kT$ is equal to 1.61×10^{-20} V · C at room temperature

The thermal noise is proportional to the frequency bandwidth Δf. It can be represented by the voltage source in series with resistor R, or by the current source in parallel to the resistor R. The maximum noise

power can be delivered to the load when $R_L = R$. In this case maximum noise power in the load is $kT\Delta f$. The noise power density $dP_n/df = kT$, and it is independent of frequency. Thus, the thermal noise is the white noise. The RMS noise voltage and the RMS noise current are proportional to the square root of the frequency bandwidth Δf. The thermal noise is associated with every physical resistor in the circuit. In a bipolar transistor, the thermal noise is generated mainly by series base, emitter, and collector resistances.

Spectral density of the equivalent voltage and currant thermal noise are given by

$$S_{vR} = 4kT_kR \tag{3.76}$$

or

$$S_{iG} = 4kT_kG \tag{3.77}$$

These spectral noise densities are constant up to 1 THz and it is proportional to temperature and to resistance of elements and as such can be used to indirectly measure:

- The device temperature
- Series distributed resistances of bipolar transistors (primarily base resistance)
- Quality of contacts and connections

3.2.2 Shot Noise

Shot noise is associated with a discrete structure of electricity and the individual carrier injection through the *pn* junction. In each forward biased junction, there is a potential barrier which can be overcome by the carriers with higher thermal energy. This is a random process and the noise current is given by

$$\overline{i_n^2} = 2qI\Delta f \tag{3.78}$$

Spectral density of the shot noise is temperature independent and it is proportional to the junction current:

$$S_{is} = 2qI \tag{3.79}$$

where

q is the electron charge
I is the forward junction current

Shot noise is usually considered as a current source connected in parallel to the small-signal junction resistance. The measurement of shout noise in modern nanoscale devices is relatively difficult since measured values of current are in the range of 100 fA.

Shot noise has to be proportional to the current and any deviation from this relation can be used to evaluate parasitic leaking resistances. It can be used for diagnosis of photodiodes, Zener diodes, avalanche diodes, and Schottky diodes.

3.2.3 Generation–Recombination Noise

The generation–recombination noise is caused by the fluctuation of number of carriers due to existence of the generation–recombination centers. Variation of number of carriers leads to changes of device conductance. This type of noise is function of both temperature and biasing conditions. The spectral density of the generation–recombination noise is described by

$$\frac{S_{g-r}(f)}{N^2} = \frac{(\overline{\Delta N})^2}{N^2} \cdot \frac{4\tau}{1 + (2\pi f \cdot \tau)^2} \tag{3.80}$$

where
$(\overline{\Delta N})^2$ is the variance of the number of carriers N
τ is the carrier lifetime

Spectral density is constant up to the frequency $f_{g-r} = 1/(2\pi\tau)$, and after that is decreasing proportionally to $1/f^2$.

In the case when there are several types of generation–recombination centers with different carrier life time the resultant noise spectrum will be a superposition of several distributions described by Equation 3.80. Therefore the spectral distribution of noise can be used to investigate various generation–recombination centers. This is an alternative method to deep level transient spectroscopy (DLTS) to study generation–recombination processes in semiconductor devices.

3.2.4 *1/f* Noise

The *1/f* noise is the dominant noise in the low-frequency range and its spectral density is proportional to *1/f*. This noise is present in all semiconductor devices under biasing. This noise is usually associated with material failures or with imperfection of a fabrication process. Most of research results conclude that this noise exists even for very low frequencies up to 10^{-6} Hz (frequency period of several weeks). This noise is sometimes used to model fluctuation of device parameters with time. There are two major models of *1/f* noise:

- Surface model developed by McWhorter in 1957 [7]
- Bulk model developed by Hooge in 1969 [8]

The simplest way to obtain $1/f$ characteristics is to superpose many different spectra of generation–recombination noises, where free carriers are randomly trapped and released by centers with different life times. This was the basic concept behind McWhorter model where it was assumed that

- In the silicon oxide near the silicon surface there are uniformly distributed trap centers
- Probability of the carrier penetration to trap centers is decreasing exponentially with the distance from the surface.
- Time constants of trap centers increases with the distance from the surface
- Trapping mechanisms by separate centers are independent

The resulted noise spectral density is given by

$$S_{1/f} \propto (\overline{\Delta N})^2 \int_{\tau_1}^{\tau_2} \frac{1}{\tau} \frac{4\tau}{1 + \omega\tau^2} \cdot d\tau = (\overline{\Delta N})^2 \cdot \frac{1}{f} \quad \text{for } 1/\tau_2 \ll \omega \ll 1/\tau_1 \tag{3.81}$$

The spectral density is constant up to frequency $f_2 = 1/(2\pi\tau_2)$, then is proportional to $1/f$ between f_2 and $f_1 = 1/(2\pi\tau_1)$, from frequency f_1 is proportional to $1/f^2$. The McWhorter model is primarily used for MOS devices.

For bipolar transistor Hooge bulk model is more adequate. In this noise model Hooge uses in the carrier transport two scattering mechanisms of carries: scattering on the silicon lattice and scattering on impurities. He assumed that only scattering on the crystal lattice is the source of the *1/f* noise, while scattering on the impurities has no effect on noise level. All imperfections of the crystal lattice leads to large *1/f* noise.

The noise spectral density for the Hooge model is

$$S_{1/f} = \frac{\alpha_H \cdot I^{\beta}}{f^{\gamma} \cdot N} \tag{3.82}$$

where
$\alpha_H = 2 \cdot 10^{-3}$ is the Hooge constant [8]
β and γ are material constants
N is the number of carriers

Later [9] Hooge proposed to use α_H as variable parameter, which in the case of silicon devices may vary from $5 \cdot 10^{-6}$ to $2 \cdot 10^{-3}$.

The $1/f$ noise is increasing with the reduction of device dimensions and as such is becoming a real problem for devices fabricated in nanoscale. The level of *1/f* noise is often used as the measure of the quality of devices and its reliability. Devices fabricated with well-developed technologies usually have much smaller level of *1/f* noise. The *1/f* noise (flicker noise) sometimes is considered to be responsible for the long term device parameter fluctuation.

3.2.5 Noise $1/f^2$

The noise $1/f^2$ is a derivative of *1/f* noise and it is observed mainly in metal interconnections of integrated circuits. It has become more evident for very narrow connections where there is a possibility of electromigration due to high current densities. In aluminum the electromigration begins at current densities of 200 $\mu A/\mu m^2$ and noise characteristics changes from $1/f^2$ to $1/f^{\gamma}$, where $\gamma > 2$. Also the noise level increases proportionally to the 3rd power of the biasing current:

$$S_{1/f^2}(f) = \frac{C \cdot J^{\beta}}{f^{\gamma} \cdot T} \cdot \exp(-E_a/k \cdot T) \tag{3.83}$$

where
$\beta \geq 3$, $\gamma \geq 2$
C is experimentally found constant
E_a activation energy of the electromigration
$k = 8.62 \cdot 10^{-5}$ eV/K is the Boltzmann constant

The degeneration of metallic layer is described by

$$\nu_d \propto J^n \exp(-E_a/k \cdot T) \tag{3.84}$$

Since Equations 3.83 and 3.84 have a similar character therefore the $1/f^2$ noise can be used as the measure of the quality of metal interconnections. This is a relatively fast and accurate method to estimate reliability of metal interconnections.

3.2.6 Burst Noise—RTS Noise

The burst noise is another type of noise at low frequencies [3,4]. Recently this noise is described as RTS noise. With given biasing condition of a device the magnitude of pulses is constant, but the switching time is random. The burst noise looks, on an oscilloscope, like a square wave with the constant magnitude, but with random pulse widths (see Figure 3.64). In some cases the burst noise may have not two but several different levels.

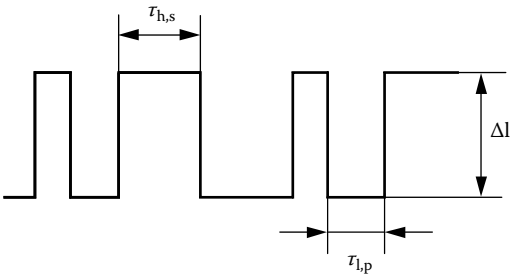


FIGURE 3.64 Example of RTS noise waveform.

Spectral density of the RTS noise has similar form like generation–recombination noise:

$$S_{RTS} = C\frac{4 \cdot (\Delta I)^2}{1 + (2\pi f / f_{RTS})^2} \tag{3.85}$$

where

$$C = \frac{1}{(\overline{\tau_l} + \overline{\tau_h}) \cdot f_{RTS}^2}$$

$f_{RTS} = \frac{1}{\tau} = \frac{1}{\tau_l} + \frac{1}{\tau_h} = \frac{\overline{\tau_l} + \overline{\tau_h}}{\overline{\tau_h} \cdot \overline{\tau_l}}$ is the corner frequency, below this frequency spectrum of the RTS noise is flat
$\overline{\tau_l}$ is the average time of pulses at low level
$\overline{\tau_h}$ is the average time of pulses at high level

$$\overline{\tau_l} = \frac{1}{P}\sum_{i=1}^{P} \tau_{l,p}$$

$$\overline{\tau_h} = \frac{1}{S}\sum_{j=1}^{S} \tau_{h,s}$$

The intensity of the RTS noise depends on the location of the trap center with the reference to the Fermi level. Only centers in the vicinity of Fermi levels are generating the RTS noise. These trapping centers, which are a source for RTS noise, are usually the result of silicon contamination with heavy metals or lattice structure imperfections.

In the SPICE program the burst noise is often approximated by

$$\overline{i_n^2} = K_B \frac{I_D^{A_B}}{1 + \left(\frac{f}{f_{RTS}}\right)^2} \Delta f \tag{3.86}$$

where K_B, A_B, and f_{RTS} are experimentally chosen parameters, which usually vary from one device to another. Furthermore, a few different sources of the burst noise can exist in a single transistor. In such a case, each noise source should be modeled by separate Equation 3.85 with different parameters (usually different corner frequency f_{RTS})

Kleinpenning [10] showed that RTS noise exists with devices with small number of carriers, where a single electron can be captured by a single trapping center. RTS noise is present in submicrometer MOS transistors and in bipolar transistors with defected crystal lattice. It is present in modern SiGe transistors.

This noise has significant effect at low frequencies. It is function of temperature, collector current, induced mechanical stress, and also radiation. In audio amplifiers the burst noise sounds as random shoots, which are similar to the sound associated with making popcorn. Obviously, bipolar transistors with large burst noise must not be used in audio amplifiers and in other analog circuitry. The burst noise was often observed in epiplanar bipolar transistors with large β coefficients. It is now assumed that devices fabricated with well developed and established technologies do not generate the RTS noise. This is unfortunately not true for modern nanotransistors and devices fabricated with other than silicon materials.

3.2.7 Avalanche Noise

The avalanche noise is another noise component, which can be found in bipolar transistors. For large reverse voltages on the collector junction, the collector current can be multiplied by the avalanche

phenomenon. Carriers in the collector–base junctions gain energies in high electrical field, then lose this energy during collision with the crystal lattice. If the energy gained between collisions is large enough, then during collision another pair of carriers (electron and hole) can be generated. This way the collector current can be multiplied. This is a random process and obviously the noise source is associated with the avalanche carrier generation. The magnitude of the avalanche noise is usually much larger than any other noise component. Fortunately, the avalanche noise exists only in the *pn* junction biased with a voltage close to the breakdown voltage. The avalanche phenomenon is often used to build the noise sources [5].

Spectral density of the avalanche noise is frequency independent

$$S_{\mathrm{I}}(f) = \frac{2qI}{(2\pi f \cdot \tau)^2} \tag{3.87}$$

where I is an average value of the reverse biasing current.

3.2.8 Noise Characterization

Many different methods are used in literature for noise characterization. Sometimes the noise is characterized by an equivalent noise resistance, sometimes by an equivalent noise temperature, sometimes by an equivalent RMS noise voltage or current or sometimes by a noise figure.

3.2.8.1 Equivalent Noise Voltage and Current

The equivalent noise voltage or current is the most commonly used method for modeling the noise in semiconductor devices. The equivalent diagram of the bipolar transistor, including various noise components, is shown in Figure 3.65. The noise components are given by

$$\overline{i_{\mathrm{B}}^2} = \frac{4kT\Delta f}{r_{\mathrm{B}}}, \quad \overline{i_{\mathrm{E}}^2} = \frac{4kT\Delta f}{r_{\mathrm{E}}}, \quad \text{and} \quad \overline{i_{\mathrm{C}}^2} = \frac{4kT\Delta f}{r_{\mathrm{C}}} \tag{3.88}$$

$$\overline{i_{\mathrm{C}}^2} = 2qI_{\mathrm{C}}\Delta f \tag{3.89}$$

$$\overline{i_{\mathrm{B}}^2} = 2qI_{\mathrm{B}}\Delta f + K_{\mathrm{F}}\frac{I_{\mathrm{B}}^{A_{\mathrm{F}}}}{f}\Delta f + K_{\mathrm{B}}\frac{I_{\mathrm{B}}^{A_{\mathrm{B}}}}{1+(f/f_{\mathrm{B}})^2}\Delta f \tag{3.90}$$

Thermal noise is associated with physical resistors only, such as base, emitter, and collector series resistances. The small-signal equivalent resistances, such as r_π and r_o, do not exhibit thermal noise.

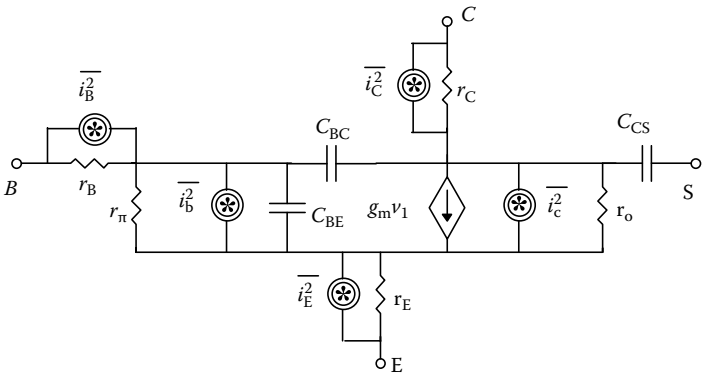


FIGURE 3.65 Equivalent diagram of the bipolar transistor which includes noise sources.

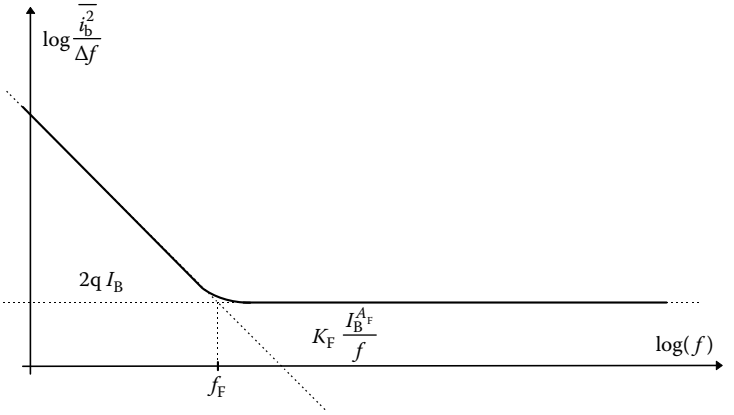


FIGURE 3.66 Bipolar transistor noise as a function of frequency.

The shot noise is associated with both collector and base currents. It was found experimentally that the $1/f$ noise and the burst noise are associated with the base current. The typical noise characteristic of a bipolar transistor is shown in Figure 3.66. The corner frequency of the $1/f$ noise can vary form 10 Hz to 1 MHz.

3.2.8.2 Equivalent Noise Resistance and Noise Temperature

The noise property of a two-port element can be described by a noise current source connected in parallel to the output terminals as Figure 3.67a shows. Knowing that noise current can be expressed as the shot noise of the DC device current the two-port noise can be expressed by means of an equivalent DC noise current

$$I_{eq} = \frac{\overline{i_n^2}}{2q\Delta f} \tag{3.91}$$

Another way to model the two-port noise in the two-port is to use the thermal noise at the input. This can be done using an additional "noisy" resistor connected to the input as Figure 3.67b shows

$$R_n = \frac{\overline{v_{n1}^2}}{4kT\Delta f} = \frac{\overline{v_{n2}^2}}{A_V^2 4kT\Delta f} \tag{3.92}$$

where

A_V is the voltage gain of the two-port

v_{n1}^2 and v_{n2}^2 are equivalent noise voltage sources at the input and the output, respectively

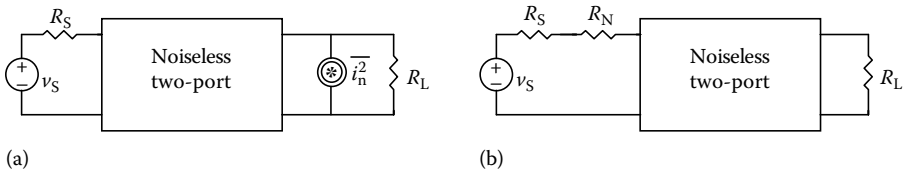


FIGURE 3.67 Noise characterization for two-ports, (a) using the noise source at the output, (b) using noise resistance R_n at the input.

The equivalent noise resistance is not a very convenient way to represent the noise property of the two-port. This additional resistance R_n must not be on the circuit diagram for small-signal analysis. To overcome this difficulty the concept of the equivalent noise temperature was introduced. This is a temperature increment of the source resistance required to obtain the same noise magnitude at the output if this source resistance is the only noise source. The noise temperature can be calculated from the simple formula

$$T_n = \frac{R_n}{R_s} 290^\circ \text{ K} \tag{3.93}$$

where R_n and R_s are shown in Figure 3.67b. It is customary to use 290° K as the reference room temperature for the noise temperature calculations.

3.2.8.3 Noise Figure

The noise figure is the ratio of the output noise of the actual two-port to the output noise of the ideal noiseless two-port when the resistance of the signal source R_s is the only noise source.

$$F = 10 \log \left(\frac{\text{total output noise}}{\text{output noise due to the source resistance}} \right) \tag{3.94}$$

The noise figure F is related to the noise resistance and the noise temperature in the following way

$$F = 10 \log \left(1 + \frac{R_n}{R_s} \right) = 10 \log \left(1 + \frac{T_n}{290^\circ \text{K}} \right) \tag{3.95}$$

The noise figure F is the most common method of noise characterization.

References

1. A. Van der Ziel, *Noise*. Prentice-Hall, New York, 1954.
2. J. L. Plumb and E. R. Chenette, Flicker noise in transistors, *IEEE Transactions on Electronic Devices*, ED-10, 304–308, Sept. 1963.
3. R. C. Jaeger and A. J. Broderson, Low-frequency noise sources in bipolar junction transistors, *IEEE Transactions on Electron Devices*, ED-17, 128–134, Feb. 1970.
4. R. G. Meyer, L. Nagel, and S. K. Lui, Computer simulation of $1/f$ noise performance of electronic circuits, *IEEE Journal of Solid State Circuits*, SC-8, 237–240, June 1973.
5. R. H. Haitz, Controlled noise generation with avalanche diodes, *IEEE Transactions on Electron Devices*, ED-12, 198–207, April 1965.
6. P. R. Gray, R. G. Mayer, *Analysis and Design of Analog Integrated Circuits*, 3rd ed. John Wiley & Sons, New York, 1993.
7. A. L. McWhorter, $1/f$ noise and germanium surface prosperities. *Semiconductor Surface Physics*, Ed. R. H. Kingdton. University of Pennsylvania Press, Philadelphia, PA, 1957, pp. 207–228.
8. F. N. Hooge, $1/f$ noise is no sourface effect. *Physics Letters*, 29A(3), 139–140, 1969.
9. F. N. Hooge, The relation between $1/f$ noise and number of electrons. *Physica B*, 162, 334–352, 1990.
10. T. G. M. Kleinpenning, On $1/f$ noise and random telegraph noise in very small electronic devices. *Physica*, B164, 331–334, 1990.

4

RF Communication Circuits

4.1 Introduction 4-1
4.2 System Level RF Design 4-2
General Overview • RF System Performance Metrics • RF Transceiver Architectures
4.3 Technology 4-9
Active Devices • Passive Devices
4.4 Receiver 4-12
LNA • Down Converter
4.5 Synthesizer 4-16
Topology • Oscillator • Prescaler • Fractional-*N* Synthesis
4.6 Transmitter 4-20
Up versus Down Conversion • CMOS Mixer Topologies • Power Amplifier
References 4-29

Michiel Steyaert
Catholic University of Leuven

Wouter De Cock
Catholic University of Leuven

Patrick Reynaert
Catholic University of Leuven

4.1 Introduction

During the last decade of last century, the world of wireless communications started to grow rapidly. Today, cellular handsets are the largest consumer market in the world. The main trigger was the introduction of digital coding and digital signal processing in wireless communications. The aggressive scaling of CMOS process technology driven by the memory and microprocessor market made CMOS a logical choice for integration of digital signal processing in wireless applications. The development of these high performance, low-cost CMOS technologies allowed integration of enormous amount of digital functionality on one chip. This enabled the use of sophisticated modulation schemes, complex demodulation algorithms, and high-quality error detection and correction to produce high data rate communication channels bringing the Shannon limit in sight [1].

The radio frequency (RF) front-ends are the interface between the antenna and the digital modem of the wireless transceiver. They have to detect very weak signals (μV) that come in at a very high frequency (10 GHz), and at the same time transmit high-power levels (up to several watts) at the same high frequencies. This requires high-performance analog circuits, like filters, amplifiers, and mixers that translate the incoming modulated data between the antenna and the A/D conversion and digital signal processing. Consumer electronic markets are mainly driven by low-cost and low-power consumption. This makes the RF front-ends the bottleneck for future wireless applications. Low-cost and low-power are both linked to high integration level. A high level of integration renders a significant space, cost, weight, and power reduction. A higher degree of integration requires less discrete components reducing the bill of materials cost. Keeping signals on chip greatly reduces power consumption since less I/O drivers are needed. Many

different techniques to obtain a higher degree of integration have been presented over the years [2–5]. This chapter introduces and analyzes some advantages and disadvantages and their fundamental limitations.

Parallel to the trend for further integration, there is the trend to integrate RF circuitry in CMOS technologies. While digital baseband processing has already been implemented in CMOS technology in several product generations, CMOS RF has only recently made its pace forward. For long time, many design houses believed complicated Mixed-Signal RF CMOS chips were impossible to realize. The main objective against CMOS RF was the lack of high-Q passive components and its poor noise performance. It took the persistence of some academic institutions and some pioneering firms to prove them wrong. It is clear that RF CMOS full potential would not have been unfold if only stand-alone radios were developed. CMOS RF systems on chip today implement all radio building blocks including phase-locked-loop (PLL), low-noise amplifier (LNA), power amplifier (PA), up- and down-conversion mixers, filters, and antenna switch. Furthermore, they include all digital baseband processing circuitry and ROM memory [6,7]. This reveals the real strength of CMOS RF over other ''better-suited'' technologies like Si Bipolar, BiCMOS, and Silicon Germanium (SiGe). Putting together RF and baseband in one chip permits compensation of lower radio performance with less expensive digital signal processing circuits, making its performance competitive with SiGe radios. Together with a possible 75% reduction of discrete components, RF CMOS offers the cheapest solution if one pursues the ultimate goal: A single chip including the physical layer (PHY) as well as the media access control (MAC) together with an MAC processor, memory, and I/O such as USBports or peripheral component interconnect (PCI) interfaces.

RF CMOS is not a matter of just replacing bipolar transistors with their CMOS counterpart. It requires a whole range a new architectures, techniques, and a high integration level. When compared with CMOS, SiGe requires less power for a certain gain and achieves a lower noise figure. The biggest drawback of CMOS is its inferior 1/f noise performance. This will only increase with the introduction of high-K dielectric materials in the gate of future CMOS technology nodes. CMOS design engineers therefore went looking for new topologies to reduce the impact of 1/f noise on the radio performance. Another problem that had to be overcome was the lack of high-Q passive components in CMOS technology. Extra processing steps as well as innovative layout and design techniques solved this problem. First, this chapter will analyze some concepts, trends, limitations, and problems posed by technology for high frequency design. Next we will discuss a variety of architectures used in modern RF CMOS transceivers. In the rest of the chapter, we will take a closer look at the different building blocks that appear in a typical RF transceiver. We will split this up between down-conversion, up-conversion, and frequency synthesis.

In the final section, we will take a look at RF CMOS's last barrier: RF power transmission. As CMOS gate lengths shrink, lower voltages are tolerated at the transistor terminals. High-quality impedance converters must therefore be placed between the antenna and the transistor's drain for high power transmissions. These are not available yet in integrated form. One of the major bottlenecks in CMOS PAs is combining high efficiency with high linearity. For high power transmission, designers are obliged to bias the PA high in its saturation region where linearity is low. Therefore, todays integrated PAs are limited to constant envelope modulation schemes like GSM. High efficient PAs still remain out of reach for modulation schemes with large peak-to-average power ratios like orthogonal frequency division modulation (OFDM). This chapter discusses some circuit techniques to circumvent this bottleneck bringing the ultimate goal of a single-chip CMOS solution that is compatible with all standards and is capable of adapting itself a step closer to reality.

4.2 System Level RF Design

4.2.1 General Overview

One of the main challenges facing the RF design engineer originates from the transmission medium used by RF systems. RF systems communicate through AIR by means of electromagnetic waves. Using air as transmission medium has one huge advantage: it gives the transceiver the ability to be mobile. However, there are some disadvantages to this high degree of freedom. There exists only one medium air, which is

consequently used by numerous applications. An overview of these applications and the part of the spectrum they use can be found on the Website of the National Telecommunications and Information Administration (NTIA) [8]. As a result, RF systems operate in a filled spectrum. Receivers will not only detect the wanted signals own to the application, but will also pick up other signals that will consequently be amplified and detected. These unwanted detected signals are called interferes. If the interferer is sufficiently large, it can corrupt the wanted signals preventing them to be properly demodulated and understood. On the transmit side of the application, unwanted signals are generated and transmitted. They are picked up by other applications and can distort their performance. These unwanted transmitted signals are called spurious signals. It is the designer's responsibility to keep these interferers and spurious signals as low as possible. Based on the earlier discussion, it is clear that one needs a regulator to manage this spectrum use. In the United States, this is done using a dual organizational structure; NTIA manages the federal government's use of the spectrum while the Federal Communications Commission (FCC) [9] manages all other uses.

Signals traveling through air also suffer from attenuation. There are several mechanisms causing attenuation such as free-space dispersion, fading, and multipath. These mechanisms depend heavily on the distance between transmitter and receiver, the frequency of transmission, and the environment. Discussion of these mechanisms, however, is beyond the scope of this text. More information concerning these topics can be found in Refs. [10,11]. As a result of these mechanisms, one can expect the received signal power to have a large variation since the distance between transmitter and receiver can change considerably due to the mobility. Performance of RF communication systems is also degraded by thermal noise. Noise is, like in other communication systems, the limiting factor when dealing with weak signals. The noise energy consists of two contributors. First, there is thermal noise which is determined by temperature and bandwidth and is out of control of the designer. On the other hand, there is system noise. This kind of noise can, within limits, be controlled by the designer to allow a certain minimum level of signal power to be detected by the system.

In the next sections, we will take a closer look at the challenges described in the earlier discussion. First, we will take a brief look at the tools and metrics RF designers use to describe and control the performance of their system in the presence of interferers and noise. We will end this section with a discussion of some commonly used transceiver architectures.

4.2.2 RF System Performance Metrics

As described in Section 4.2.1, the lowest signal power level that can be detected correctly by a receiver is limited by noise. The lowest power level that can be detected is usually called the receiver sensitivity. The receiver sensitivity is related to the signal-to-noise ratio (SNR) at the end of the receive chain (baseband). The SNR at baseband is determined by the bit error rate (BER) required by the application. It is usually expressed in terms of E_b/N_o. E_b is the energy per received bit and N_o is the noise power density received together with the bit. The relation between E_b/N_o and BER depends on the modulation scheme used in the application (e.g., Gaussian minimum shift keying (GMSK) in global system for mobile communications (GSM)) and is beyond the scope of this text. More information can be found in Ref. [12]. The SNR can be expressed in function of E_b/N_o as follows:

$$\mathrm{SNR} = \frac{S}{N} = \frac{E_b}{N_o} \times \frac{f_b}{B} \tag{4.1}$$

where

f_b is the bit rate

B is the receiver noise bandwidth

Note that the overall system noise at baseband N is the sum of thermal and circuit noise. This leads to a figure of merit that describes the circuit's performance. It is called noise figure when expressed in decibels and noise factor otherwise. Noise factor or figure is a measure for the excess noise that is contributed by

the circuit to the overall noise and is defined as the ratio between the SNR at the input of the receiver (SNR_i) and the SNR at the output of the receiver (SNR_o)

$$NF = \frac{\mathrm{SNR_i}}{\mathrm{SNR_o}} = \frac{(S/N)_\mathrm{i}}{(S/N)_\mathrm{o}} \tag{4.2}$$

If the receiver consists of different building blocks, one may want to know the noise figures of the different blocks and not only the overall noise figure. One can prove that in case of a series connection [12]

$$\mathrm{NF_{total}} = \mathrm{NF_1} + \frac{\mathrm{NF_2} - 1}{G_1} + \frac{\mathrm{NF_3} - 1}{G_1 G_2} + \cdots \tag{4.3}$$

where

NF_i are the noise factors of successive building blocks
G_i is their respective power gain

One can easily conclude from Equation 4.3 that building blocks earlier in the receive chain have a larger contribution to the overall noise figure than blocks at the end of the chain. This is the reason behind the use of an LNA at the input of an RF receiver. The large power gain combined with a low-noise figure will relax the noise specifications for the following blocks. The principle is explained in Figure 4.1. If an LNA is omitted and the mixer is put directly behind the antenna, the signal is drowned in the mixer noise and the sensitivity will be low. The power gain of the LNA, however, pushes the antenna signal above the noise floor of the mixer. As long as the output noise of the LNA is greater than the input noise of the mixer, the sensitivity is fully determined by the NF of the LNA.

RF systems often operate in an interference limited environment. Interference can also reduce receiver sensitivity. It is therefore more correct to describe the receiver sensitivity by its signal-to-noise plus interference ratio $S/(N+I)$ also known as the signal-to-noise and distortion ratio (SNDR). One of the mechanisms by which interference limits the performance is nonlinearity. It can reduce the signal power as well as increase interference. Large signals can saturate the receiver resulting in a gain compression, which reduces the signal power S. On the other hand, two large interfering signals can, due to non linearity, produce cross-product terms that fall on top of the wanted signal increasing the interference I. This cross-product generation is called intermodulation distortion (IMD). Nonlinearity performance is typically characterized by small signal linearity described by second-and third-order intercept points

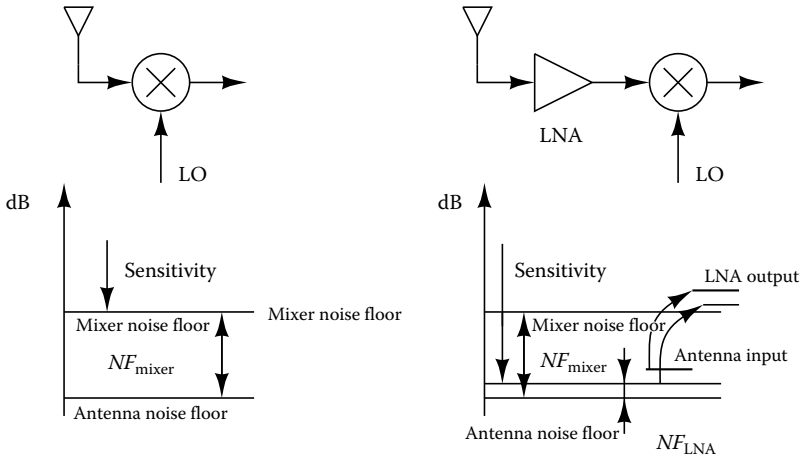


FIGURE 4.1 The benefit of using an LNA.

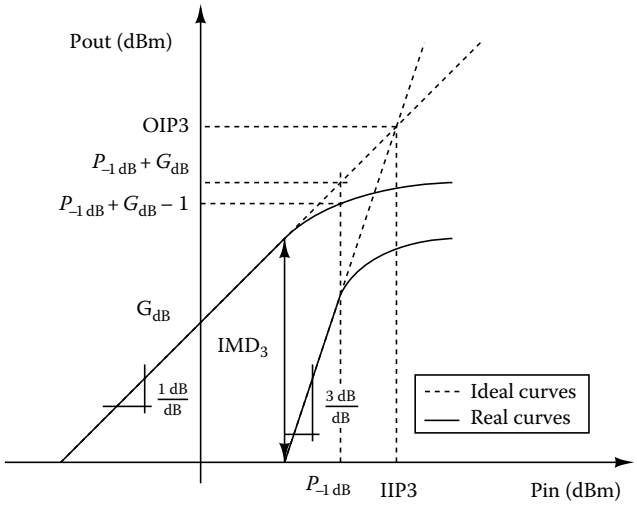


FIGURE 4.2 First- and third-order intermodulation as a function of the input power.

(IP2 and IP3) and large signal linearity described by the 1 dB compression point. Usually balanced topologies are used attenuating the second-order harmonics. Consequently, third-order nonlinearity will become the limiting factor. These concepts will be explained with the help of Figure 4.2. Gain compression is characterized by the 1 dB compression point ($P_{-1\ \mathrm{dB}}$) and is used to evaluate the ability of the system to cope with strong input or interference signals often referred to as blockers. It is defined as the input power for which the gain drops by 1 dB. By identifying the strongest signals at each stage of the design, one can calculate the required 1 dB compression point for each block in a receiver chain. As mentioned earlier, nonlinearity not only causes gain compression, but also generates IMD. This is produced by any pair of blockers that lie near the wanted signal. If two tones at f_1 and f_2 are applied to a nonlinear block, frequencies are produced not only at f_1 and f_2 but also at $2f_1 - f_2$, $2f_2 - f_1$, $3f_1$, $3f_2$, and so on. $f_1, f_2, 3f_1$, and $3f_2$ are not important since they lie far outside the frequency band of interest and can therefore be filtered out. $2f_1 - f_2$ and $2f_2 - f_1$, however, are potential problems as they can fall on top of the wanted signal band and remain unaffected by filtering. The ratio of any of the two cross products is called third-order IMD3. The output power of the intermodulation products grows at a faster rate than that of the wanted signal itself. Therefore, it follows that at a certain input power, the output power of the intermodulation signals will surpass the wanted signal. The input power level where this takes place is called the input-referred third-order intercept point (IIP3). The output power at this point is called the output-referred third-order intercept point (OIP3). Note that this is an imaginary point since gain compression occurs before this point is reached. If the receiver consists of different building blocks, one may want to know the contribution of the different building blocks to the overall linearity performance. One can prove that in case of a cascaded system

$$\frac{1}{\mathrm{IIP3}_{total}} = \frac{1}{\mathrm{IIP3}_1} + \frac{G_1}{\mathrm{IIP3}_2} + \frac{G_1 G_2}{\mathrm{IIP3}_3} + \cdots \tag{4.4}$$

where

$\mathrm{IIP3}_i$ are the input-referred third-order IPs of the successive building blocks
G_i is their respective power gain

One can conclude that, contrary to noise (see Equation 4.3), the last blocks in the receive chain has the largest influence on the overall linearity of the receiver. Equations 4.3 and 4.4 reveal a first trade-off. High gain at the input reduces noise constraints in the rest of the chain but increases the linearity requirements.

A last origin of distortion is due a nonideal local oscillator (LO) signal driving the mixers. In practice, the spectrum of an oscillator is never pure. There is always a certain amount of energy present close to the ideal LO frequency at $\omega_0 + \Delta\omega$. This can translate nearby frequency signals on top of the wanted signal also deteriorating the SNDR of the system. A figure of merit to describe this nonideal LO behavior is called the LO phase noise and is defined as the ratio of the power present in a 1 Hz band at a certain offset frequency $\Delta\omega$ from the carrier frequency ω_0 to the carrier power:

$$\mathcal{L}\{\Delta\omega\} = 10\log\left(\frac{\text{noise power in a 1 Hz band at } \omega_0 + \Delta\omega}{\text{carrier power}}\right) \tag{4.5}$$

4.2.3 RF Transceiver Architectures

In this section, a brief overview of some common transceiver structures will be discussed and contrasted to one another. The discussion will be restricted to the heterodyne transceiver, the zero-IF or direct conversion transceiver, and the low-IF transceiver. There exist numerous other types of transceivers but their properties can be understood by looking at these three structures as they are all variations or combinations of these three structures. First, the different receiver architectures will be discussed followed by there transmitter equivalent.

The heterodyne receiver has been the dominant choice in RF systems for many decades. The reason behind this is its high performance and adaptability to different standards. Figure 4.4 shows the operation of a heterodyne receiver. The broadband antenna signal is first fed to a highly selective RF filter (band select filter) that suppresses all interferers outside the wanted application band. An LNA boosts the wanted signal above the mixer noise floor and an LO generates a signal located at an offset frequency f_{IF} from the wanted signal. The result is that the following signals are down-converted by the mixer to f_{IF}

$$f_{\text{wanted}} = f_{LO} - f_{IF} \tag{4.6}$$

$$f_{\text{image}} = f_{LO} + f_{IF} \tag{4.7}$$

Not only the wanted signal is mapped onto IF (intermediate frequency), but also another signal called the image or mirror signal. This signal can corrupt the information content in such a way that the information is irreparable. To avoid this, an image reject filter is inserted before the mixer. This way, a highly attenuated version of the image signal is folded on top of the wanted signal, preventing the irreparable corruption of the information content of the signal. Figure 4.3 summarizes this operation. From Equation 4.6 and 4.7, one can see that the center of the image signal is located at a distance $2f_{IF}$ from the wanted signal. The choice of f_{IF} therefore determines the requirements for the image reject filter. If a very low f_{IF} is chosen, a very high-quality filter is needed to suppress the image frequency. To relax the filter specifications, f_{IF} is usually chosen relatively high and a series of down-conversion steps are performed. The heterodyne structure is then referred to as the superheterodyne receiver.

The heterodyne or superheterodyne receiver features a single path topology. Mismatch between different parts is not a issue here. Also LO feedthrough in the mixer is not a problem, since the wanted signal is never close to the LO frequency. In Figure 4.4, it can also be seen that the channel selection is done before the AGC-A/D structure. They will therefore only need to handle a limited dynamic range.

A drawback of the structure, however, is that all critical functions are realized with passive devices. Due to the high demands posed upon these structures, they are mostly implemented off-chip. The integratability of the heterodyne transceiver is therefore rather low. This induces an additional material cost. Moreover, the insertion loss of the passive filters needs to be compensated by a higher gain on-chip to keep the required SNR. Since the filters need to be driven at low impedance (e.g., 50 Ω), one has the choice between using complex impedance transformation structures or using low-output impedance

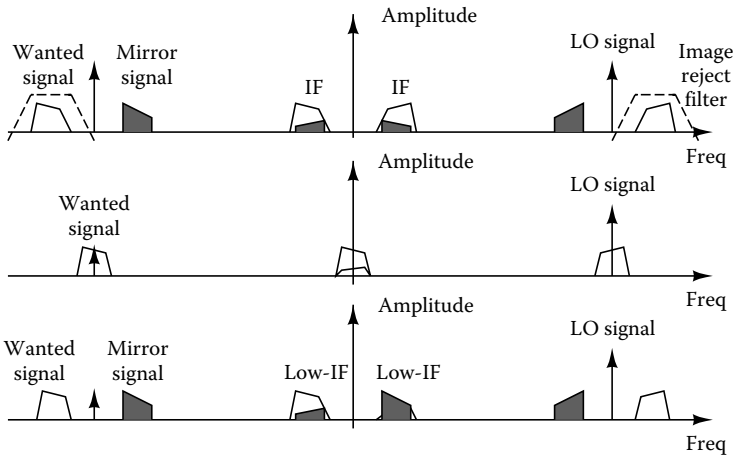


FIGURE 4.3 The down-conversion process in an IF, zero-IF, and a low-IF receiver.

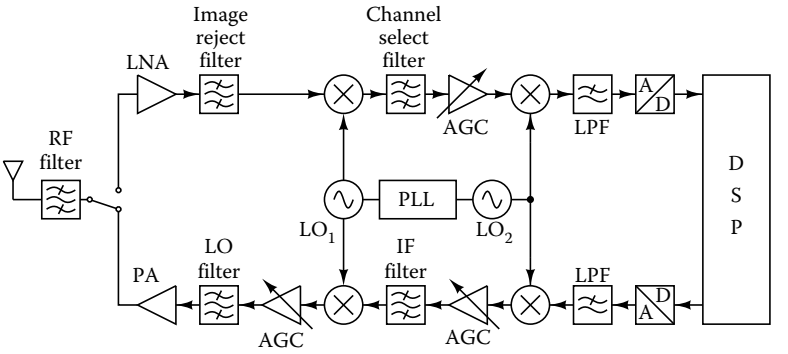


FIGURE 4.4 Heterodyne transceiver architecture.

buffers. Using low output impedance drivers, however, comes at the cost of an extra amount of extra power consumption.

The integratability however can dramatically be improved if one could find a way of getting rid of the external high-quality filters. This means looking for a way of suppressing the image frequency without filters. A first solution to this problem is obvious. Make the image signal the wanted signal or choose $f_{IF} = 0$. This solution is called the zero-IF receiver or direct-conversion receiver [13,14]. Another solution is related to the first one and is called the low-IF topology [3]. This topology takes advantage of the fact that the channels in the direct neighborhood of the wanted channel—the adjacent channels—are usually much weaker than the wanted signal and the signals laying further away. Furthermore, these frequency bands are usually regulated in the application specifications or by the FCC. So, if an IF-frequency is chosen so that the image frequency falls into this lower power bands, less image rejection is needed to keep the required SNR. Figure 4.5 shows the architecture of both a direct or zero-IF receiver and a low-IF receiver. The only difference between both can be found in the choice of IF-frequency. In a zero-IF receiver, the wanted channel is converted to DC and a mirrored version of the channel itself is superimposed onto the clean version of the signal. In a low-IF receiver, the wanted signal is down-converted to a low, nonzero IF, e.g., half the channel bandwidth, such that the mirror signal is the adjacent channel. The antenna signal is first passed through a band select filter. An LNA boosts the signal

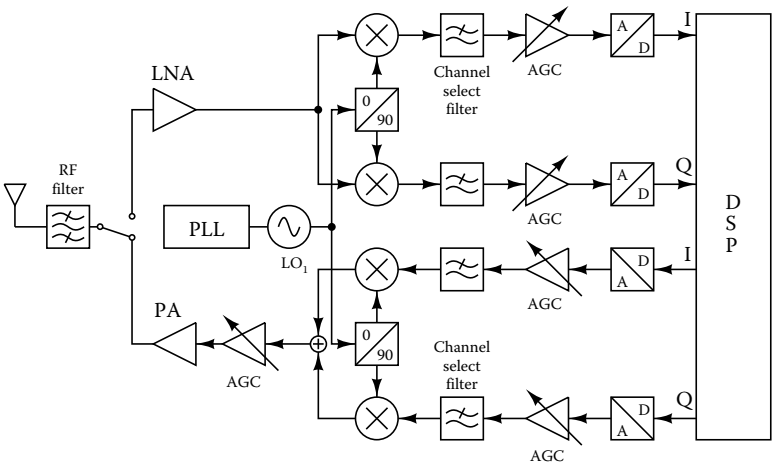


FIGURE 4.5 Direct transceiver architecture.

above the mixer noise floor. So far, there is no difference with the heterodyne receiver. After the LNA, however, the signal is fed to two different signal paths. The two signal paths are then down-converted by two mixers that are steered by two LO signals that are spaced 90° apart. The interstage filter has now become obsolete since the mirror signal will be neutralized by recombining the two signal paths after down conversion. This type of down conversion is called quadrature down conversion. Since the image signal and the wanted signal are separated in the digital signal processor (DSP), the real channel selection and image rejection are done in the digital back-end. This is a positive thing, since the digital domain is the natural biotope of CMOS. Since the image rejection and channel selection no longer rely on high-quality filtering, no external filters are required; therefore, one does not have to cope with their inevitable loss and one does not need low impedance drivers. This allows low power operation. However, the spreading of the signal over two independent signal paths has some drawbacks. The topology relies heavily on the symmetry between the two paths, every mismatch between the two paths will lead to a deterioration of the image suppression and an increased corruption of the wanted information content. Although one could think that that image rejection requirements are more relaxed for a zero-IF receiver since the image signal is a mirrored version of the wanted signal, this is not exactly true. For low-IF receivers, the image signal can be considered as noise for the wanted signal, since there is no correlation at all between the two bands. For zero-IF receivers, there is a strong correlation between image and wanted signal leading to a distortion of the wanted signal. The required image suppression is therefore dependent on the type of modulation that is used in the system. When a quadrature amplitude modulation (QAM) type modulation is used, one can calculate that the required image rejection for zero-IF is 20–25 dB while 32 dB rejection is required for low-IF systems [15]. As the wanted signals in both receivers are located at low frequencies (dc in case of zero-IF), the signal is susceptible to 1/f noise and dc-offset. Complicated feedback structures can get rid of the dc-offset; however, due to the finite time-constants in those loops, part of the signal is also canceled by the feedback. This can corrupt the signal in an unacceptable way. Low-IF topologies are less vulnerable. As long as the dc-offset does not saturate the A/D converters, there is no signal degradation. Due to the absence of filtering in the RF part, the A/D converters, however, have to deal with larger dynamic ranges. Fortunately, as the signals are at low frequencies, oversampled converters can be used which allow higher accuracies.

The same topologies exist for the transmitter side of the transceiver. The heterodyne as well as the direct up-conversion transmitter will be discussed. They are depicted in Figures 4.4 and 4.5. The early up-conversion architectures were in fact multistage architectures. They employed a number of mixing stages

and intermediate frequencies. The main advantage of this type of systems is that only one D/A converter is needed. Quadrature modulated signals are therefore generated in the digital domain. This topology puts high demands on the D/A converter since it must deliver signals at a higher IF frequency. The DSP on the other hand must be able to deliver perfectly matched I/Q signals. This approach requires the use of high-quality passives, multiple LOs. The same conclusions can be drawn as in the receiver. Due to the large number of external components, integratability is limited and power consumption will be high. Another implementation of this multistage architecture includes the use of two D/A converters. Quadrature modulated signals are then generated in the analog domain. Since they are generated at low frequencies, quadrature matching is superior. However, multiple RF filters are still needed, giving rise to a higher cost and power consumption. The topology, however, is not vulnerable to one of the main problems in the direct conversion architecture, oscillator pulling caused by the PA due to the fact that the PA output spectrum is far away from the voltage-controlled oscillator (VCO) frequency. Hereby the main problem in direct up-conversion circuits is addressed. In direct conversion transmitters, the transmitted carrier frequency is equal to the LO frequency. As can be seen in Figure 4.5, modulation and up conversion occur in the same circuit. The I/Q quadrature modulator takes the baseband (or low-IF) input signal and up-converts it directly to the desired RF frequency. This eliminates the need for RF passives and limits the number of amplifiers, mixers, and LOs. The simplicity of the architecture makes it an obvious choice when high integration levels are demanded. However, as mentioned before, the circuit suffers from one major drawback, the disturbance of the LO by the PA. This phenomenon is explained in detail in Refs. [16,17]. As the LO frequency lies in the transmit band, high demands are put on the LO/RF isolation. The system is also susceptible to I/Q mismatch errors, even the least phase mismatch or amplitude difference between I and Q path will result in distortion in the spectrum. However, the elimination of the IF stage in the transmitter leads to large saving in material cost and increases the robustness of the system as the number of discrete components that could fail is reduced. There is not only a cost saving in material cost, the direct up-converter architecture also allows a reduction in equipment size. This makes the circuit first choice for applications with stringent space constraints [18].

4.3 Technology

4.3.1 Active Devices

Since all high level or system level designs in the end need to be implemented in terms of actual active and passive components, it is no surprise that the transistor performance is of major importance for the overall system performance. It is therefore imperative to know the performance limitations of the technology one is working in and to be aware of the shortcomings of the model one is using. It is clear that conformity between measurements and simulation results will strongly depend on the accuracy of the models used with respect to the actual behavior of the devices. Although several compact models exist to describe MOSFET transistors, the BSIM [19] is considered as the de facto standard because it is the model that is generally provided by silicon foundries. Most models are quite accurate for low frequencies; however, most models fail when higher frequencies are to be modeled. ''High frequency'' means operating frequencies around 1/10th of the transistor's cutoff frequency f_t. Figure 4.6 gives an overview of f_t for different technology nodes. For a standard 0.18 μm technology with an f_t of around 50 GHz, this means 5 GHz is considered to be a high frequency. Another parameter is plotted in Figure 4.6, $f_{3\,dB}$ reflects the speed limitation of a transistor in a practical configuration. It is defined as the 3 dB point of a diode connected transistor [20] and takes into account the parasitic speed limitation due to overlap capacitances, drain-bulk junction, and gate-source capacitance while f_t only models the parasitic effect of the gate-source capacitance. In Ref. [21], an extended transistor model is presented that can be used for circuit simulation at RF frequencies. It is shown in Figure 4.7. All the extrinsic components are pulled out of the MOS transistor model, so that the MOS transistor symbol only represents the intrinsic part of the device. This allows to have access to internal nodes and model extrinsic components such as series

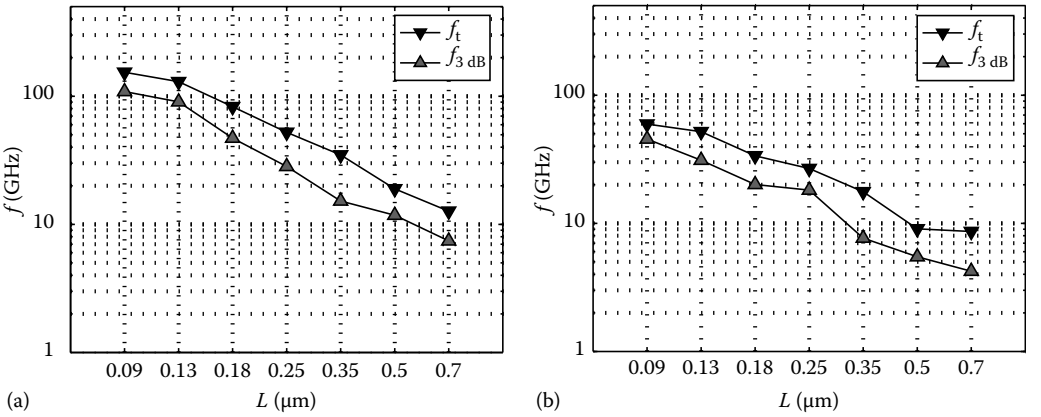


FIGURE 4.6 Maximum operating frequencies for different technology nodes.

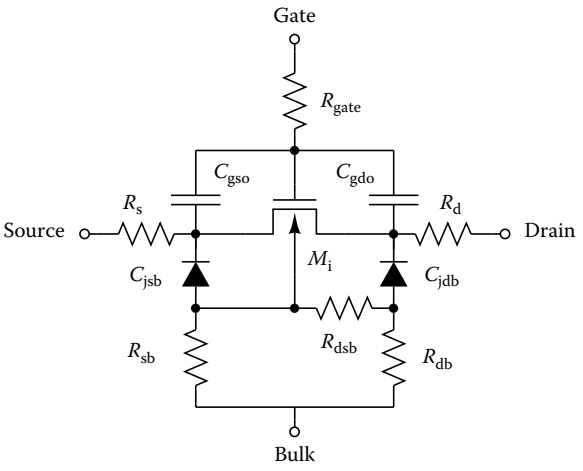


FIGURE 4.7 Extended RF transistor model.

resistances and overlap capacitances in a different way than what is available in the complete model. The source and drain series resistors are added outside the MOS model since the series resistances internal to the compact model are only used in the calculation of the I–V characteristic to account for the dc voltage drop across the source and drain. They do not add any poles and are therefore invisible for ac simulation. The gate resistance is usually not part of a MOSFET model, but plays a fundamental role in RF circuits and is therefore of outmost importance. The substrate resistors R_{dsb}, R_{sb}, and R_{db} have been added to account for the signal coupling through the substrate. Apart from the extra components added in the extended transistor model presented in Ref. [21], another point deserves some attention. The classical transistor model is based on the so-called quasi-static assumption. This means that any positive (negative) change in charge at the gate is immediately compensated by a negative (positive) change of charge in the channel. In reality, however, there will always be a delay in the charge buildup in the channel. Individual electrons (holes) will need a finite time to travel from bulk to the channel. This effect is called the non-quasi-static effect and has been described in Ref. [22–24]. This effect can be modeled by adding a resistance in series with the gate-source capacitance, introducing an extra time constant in the model.

$$\tau_{gs} = \frac{C_{gs}}{5g_m} = \frac{1}{5\omega_t} \tag{4.8}$$

This model is valid in strong innversion and within the long channel approximation. Although one could think that this effect is neglectable at realistic operating frequencies much lower than f_t, in bandpass applications, the gate-source capacitance can be tuned away by an inductor making the input impedance of the transistor purely resistive.

4.3.2 Passive Devices

For a long time, CMOS RF integration was believed to be impossible due to the poor quality of passive devices. Smaller CMOS geometries and innovative design and layout [25–27], however, have enabled high-quality passive components at high frequency to be integrated on chip. Four passive devices (resistors, inductors, capacitors, and varactors) will be discussed. First, one needs of figure of merit to qualify these passive devices. In general, the Q-factor is used for this purpose. Although there exist several definitions for the Q-factor, the most fundamental definition is based on the ratio between the maximum energy storage and the average power dissipation during one cycle in the device.

$$Q = \frac{\omega W_{\max}}{P_{\text{diss}}} \tag{4.9}$$

For an overview of other definitions of the Q-factor, the reader is referred to Ref. [28]. For a purely reactive element (capacitor or inductor), current through the element and voltage over the element are 90° out of phase. Hence, no power is dissipated in it. In real live, however, a certain amount of power will always be dissipated. Power dissipation supposes the presence of a resistance and a resistance always generates thermal noise. The Q-factor consequently is also a way of describing the pureness of a reactive device. Figure 4.8 shows some very common structures used in the modeling of reactive components used in RF circuits together with their Q-factor according to Equation 56.9. Low-Ohmic resistors are commonly available now in all CMOS technologies and their parasitic capacitance is such that they allow for more than high enough bandwidth. A more important passive device is the capacitor. In RF circuits, capacitors can be used for AC coupling. This enables DC-level shifting between different stages resulting in an extra degree of freedom enabling an optimal design of each stage. It also offers the possibility of lowering the power supply voltages. Another field, although not completely RF, where capacitors are commonly used is to implement switched capacitor circuits or arrays. This is favorable to using common resistors since capacitors in general offer better matching properties than resistors. The quality of an integrated capacitor is mainly determined by the ratio between the capacitance value and the value of the parasitic capacitance to the substrate. Too high a parasitic capacitor loads the transistor stages, thus reducing their bandwidth, and it causes an inherent signal loss due to a capacitive division.

The passive device, however, that got the most attention in the past is the inductor. It was long believed that high-quality integrated inductors were simply impossible in standard CMOS processes [29] and

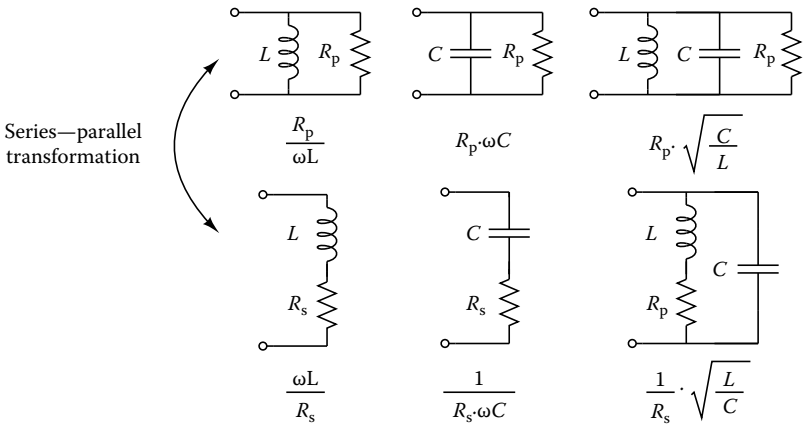


FIGURE 4.8 Quality factors of some common circuits.

could better be avoided if possible. However, due to the use of hollow spiral inductors and slightly altered process technology (thick top metal layer), one is now able to produce high-Q inductors in CMOS. The use of inductors on chip allows a further reduction of the power supply and offers compensation for parasitic capacitors by tuning them away resulting in higher operating frequencies. To be able to use integrated inductors in actual designs, an accurate model is needed. Reference [30] introduces such a model. One of the problems faced when modeling an inductor is how to model the substrate. One of the major drawbacks of inductors is the losses introduced by the substrate underneath the coil by capacitive coupling and eddy currents. This reduces the quality factor of the inductor.

A last passive component that is often encountered in RF CMOS designs is the varactor. It is mostly used for implementing tunable RF filters and VCOs. The different varactor types can be put in two classes: junctions and MOS capacitors. The latter can be used in accumulation and in inversion mode. For all cases, the devices have to be placed in a separate well to be able to use the well potential as the tuning voltage. For a standard NWELL process, the available configurations are therefore limited to p^+/n^- junction diodes and PMOS capacitors. When comparing the different varactor types, one should look at the following specifications: the varactor should offer a high Q-factor, the tuning range over which the capacitance can be varied should be compatible with the supply voltages used in the design, the physical structure should be as compact as possible to limit the area and its capacitance variation should be uniform over the complete tuning range as this makes feedback design easier. For an extended discussion about the different varactor types and their performance, the reader is referred to Ref. [27].

4.4 Receiver

4.4.1 LNA

The importance of the LNA has been explained earlier. The LNA is used to boost the received signal above the mixer noise floor. It is therefore critical that the LNA itself produces little noise. The noise figure of an LNA embedded in a 50 Ω system is defined as

$$\mathrm{NF} = 10 \log_{10} \left(\frac{\text{LNA output noise}}{\text{LNA output noise if the LNA itself was noiseless}} \right) \tag{4.10}$$

that is the real output noise power (dv^2/Hz) of the LNA (consisting of the amplified input noise power and all noise contributions generated in the LNA itself) divided by the amplified input power. Figure 4.9 shows some common input structures. Figure 4.9a shows a nonterminated common source input stage. Figure 4.9b shows the same input stage but now with an impedance matching at the input. Figure 4.9c shows the common gate input structure and finally Figure 4.9d shows a transimpedance amplifier

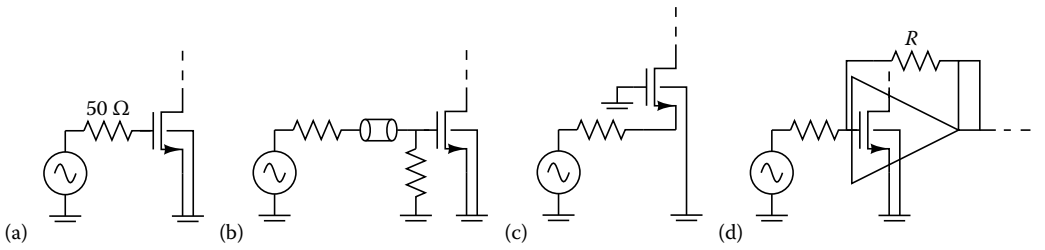


FIGURE 4.9 Some common LNA topologies.

structure that is commonly used for wideband applications. Their respective noise figures can be approximated with the following equations:

$$\text{Common source nonterminated (Figure 4.9a):} \quad \text{NF} = 1 + \frac{1}{50 \cdot g_m} \tag{4.11}$$

$$\text{Common source terminated (Figure 4.9b):} \quad \text{NF} = 2 + \frac{1}{50 \cdot g_m} \tag{4.12}$$

$$\text{Common gate (non)terminated (Figure 4.9c):} \quad \text{NF} = \left[\frac{1 + 50 \cdot g_m}{50 \cdot g_m}\right]^2 + \frac{1}{50 \cdot g_m} \tag{4.13}$$

$$\text{Common source transimpedance (Figure 4.9d):} \quad \text{NF} = 1 + \frac{1}{50 \cdot g_m} \cdot \left[\frac{R + 50}{R}\right]^2 + \frac{50}{R} \tag{4.14}$$

Figure 4.10 compares the noise figures of the different topologies. It is clear that the transimpedance structure and the not terminated common source circuit are far superior compared to the other structures as far as noise is concerned. For those circuits, the NF can be approximated as

$$\text{NF} - 1 \approx \frac{1}{50 \cdot g_m} = \frac{(V_{gs} - V_T)}{2 \cdot 50 \cdot I} \tag{4.15}$$

indicating that a low noise figure requires a large transconductance in the first stage. To generate this transconductance with high power efficiency, we need to bias the transistor in the region with a large transconductance efficiency, i.e., low $V_{gs} - V_T$. This, however, will result in a large gate-source capacitance limiting the bandwidth of the circuit. Together with the 50 Ω source resistance, the achievable bandwidth is limited by

$$f_{3\ \text{dB}} = \frac{1}{2\pi 50 C_{gs}} \tag{4.16}$$

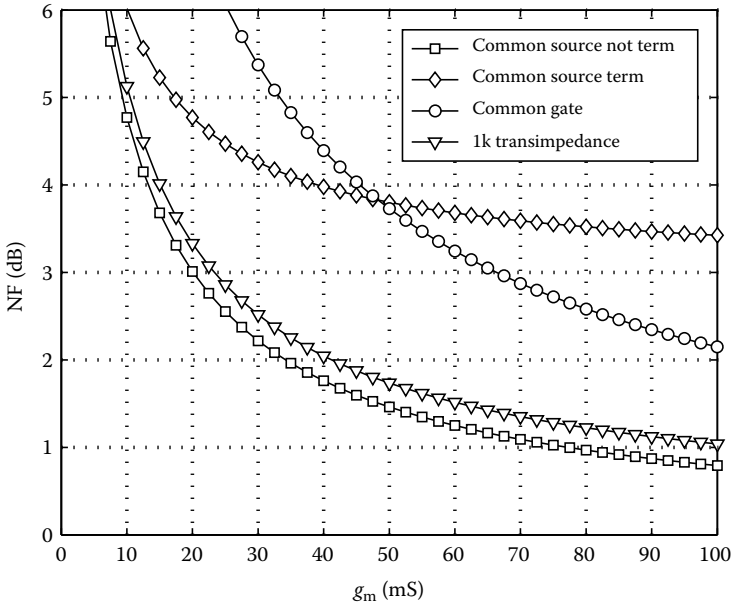


FIGURE 4.10 LNA input structure performance comparison.

When using the well-known approximative expression for the cutoff frequency of a transistor f_T

$$f_T = \frac{g_m}{2\pi C_{gs}}$$

one can conclude that

$$NF - 1 = \frac{f_{3\ dB}}{f_T} \tag{4.17}$$

This means that a low noise figure can only be achieved by making a large ratio between the frequency performance of a transistor, represented by f_T and the theoretical bandwidth $f_{3\ dB}$ of the circuit. Note that the $f_{3\,dB}$ used here is not the same as the one used in Section 4.3. Since f_T is proportional with $V_{gs} - V_T$, a low noise figure requires a large $V_{gs} - V_T$ and associated with it a large power drain. Only by going to deep submicron technologies will f_T become large enough to achieve low noise figures for gigahertz operation with low power consumption. In practice, the noise figure is further optimized by using noise and source impedance matching. These matching techniques often rely on inductors to cancel out parasitics by creating resonant structures. This boosts the maximum operation frequency to higher frequencies. More information concerning the design and optimization of common source LNAs can be found in Refs. [15,31].

At high antenna input powers, the signal quality mainly degrades due to in-band distortion components that are generated by third-order intermodulation in the active elements. Long channel transistors are generally described by a quadratic model. Consequently, a one transistor device ideally only suffers from second-order distortion and produces no third-order intermodulation products. As a result, high IIP3 values should easily be achieved. When transistor lengths shrink, however, third-order intermodulation becomes more important.

To start the analysis of the main mechanisms behind third-order intermodulation, one needs an approximate transistor model. A drain current equation that is strongly related to the SPICE level 2 and level 3 model is

$$I_{ds} = \frac{\mu_0 C_{ox}}{2n} \cdot \frac{W}{L} \cdot \frac{(V_{gs} - V_T)^2}{1 + \Theta \cdot (V_{gs} - V_T)} \tag{4.18}$$

with

$$\Theta = \theta + \frac{\mu_0}{L_{eff} \cdot v_{max} \cdot n} \tag{4.19}$$

where

- θ stands for the mobility degradation due to transversal electrical fields (surface scattering at the oxide–silicon interface)
- $\mu_0/(L_{eff} \cdot v_{max} \cdot n)$ models the degradation due to longitudinal fields (electrons reaching the thermal saturation speed)

As the θ-term is small in todays technologies, it can often be neglected relative to the longitudinal term. It can be seen from Equation 4.18 that for large values of $V_{gs} - V_T$, the current becomes a linear function of $V_{gs} - V_T$. The transistor is then conducting in the velocity saturation region. For smaller values of $V_{gs} - V_T$, the effect of Θ consists apparently in linearizing the quadratic relationship, but in reality, the effect results in an intermodulation behavior that is worse than in the case of quadratic transistors. The second-order modulation will be lower, but it comes at the cost of a higher third-order intermodulation. The following equations can be found by calculating the Taylor expansions of the drain current around a certain $V_{gs} - V_T$ value [32]:

$$\text{IIP2} \cong 10 + 20\log_{10}\left(\left(V_{gs} - V_T\right) \cdot (1 + r) \cdot (2 + r)\right) \tag{4.20}$$

$$\text{IIP3} \cong 11.25 + 10\log_{10}\left(\left(V_{gs} - V_T\right) \cdot V_{sv} \cdot (1 + r)^2 \cdot (2 + r)\right) \tag{4.21}$$

where

$$V_{sv} = \frac{1}{\Theta} \tag{4.22}$$

represents the transit voltage between strong inversion and velocity saturation and

$$r = \frac{V_{gs} - V_T}{V_{sv}} \equiv \Theta \cdot \left(V_{gs} - V_T\right) \tag{4.23}$$

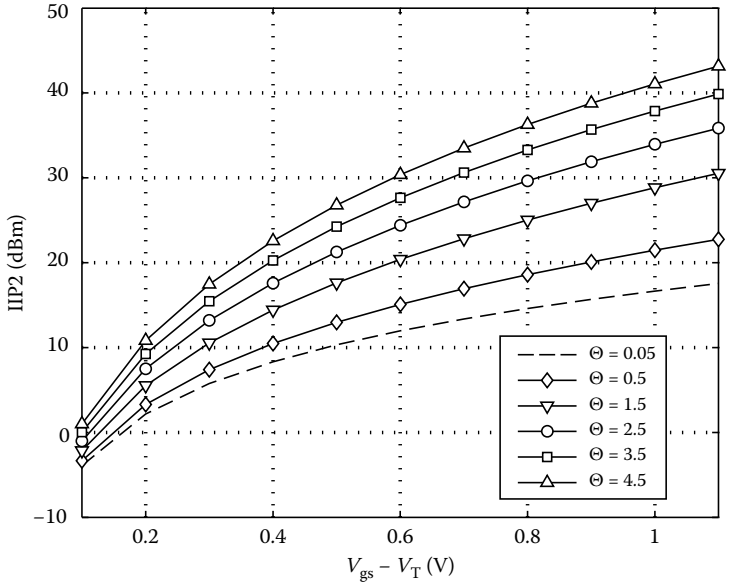
denotes the relative amount of velocity saturation. The transit voltage V_{sv} depends only on technology parameters. For deep submicron processes, this voltage becomes even smaller than 300 mV, which is very close to the $V_{gs} - V_T$ at the boundary of strong inversion. The expressions for IIP2 and IIP3 are normalized to 0 V dBm, the voltage that corresponds to a power of 0 dB in a 50 Ω resistor. For a given L_{eff}, the IIP3-value of a transistor is only a function of the gate overdrive voltage. Figure 4.11 plot the IIP2 and IIP3 in function of the gate overdrive voltage for different values of Θ. It can be seen that for a certain value of $V_{gs} - V_T$, the IIP2 increases for increasing Θ (decreasing gate lengths) which proves former statements. The picture becomes a bit more complicated when looking at the IIP3 plot. For practical values of Θ, one can distinguish two regions in the $V_{gs} - V_T$ domain. For high gate overdrive voltages, deep submicron transistors clearly exhibit better linearity because the saturation voltage becomes lower and the transistor will reach velocity saturation earlier. Short channel transistors therefore offer a maximum amount of linearity at a given power supply and require minimum $V_{gs} - V_T$ for a given IIP3. On the other hand, for low overdrive voltages, short channel transistors perform worse. Thus, to ensure a certain amount of linearity, one has to bias the transistors at a high enough overdrive voltage or apply some linearizing feedback technique (e.g., source degeneration). It can be shown that for the same equivalent g_m and the same distortion level, the required dc current is lower when local feedback is provided at the source. It comes, however, at the cost of a larger transistor and this can compromise the amplifier bandwidth.

4.4.2 Down Converter

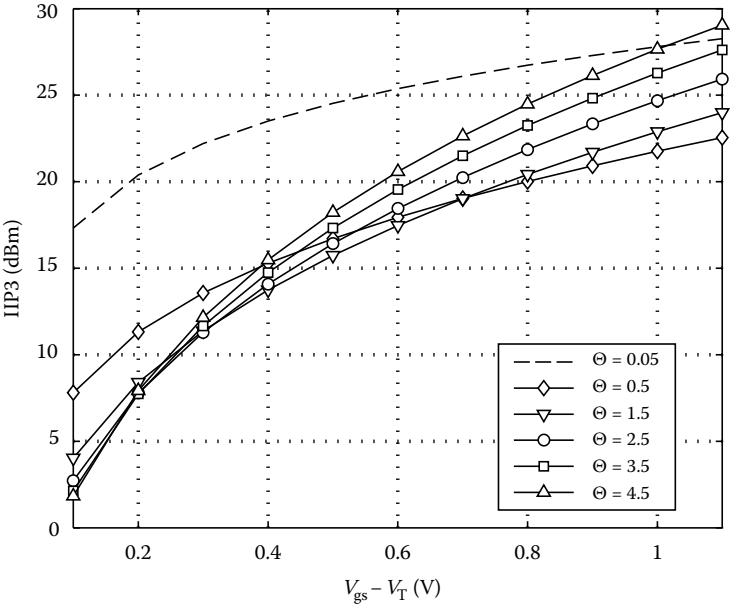
The most often used topology for a multiplier is the multiplier with cross-coupled variable transconductance differential stages. The use of this topology or related topologies (e.g., based on the square law) in CMOS is limited for high-frequency applications. Two techniques are used in CMOS: the use of the MOS transistor as a switch and the use of the MOS transistor in the linear region.

The technique often used in CMOS downconversion for its ease of implementation is subsampling on a switched-capacitor amplifier [33,34]. Here, the MOS transistor is used as a switch with a high input bandwidth. The wanted signal is commutated via these switches. Subsampling is used in order to be able to implement these structures with a low frequency op-amp. The switches and the switched capacitor circuit run at a much lower frequency (comparable to an IF frequency or even lower). The clock jitter must, however, be low so that the high frequency signals can be sampled with a high enough accuracy. The disadvantage of subsampling is that all signals and noise on multiples of the sampling frequency are folded upon the wanted signal. The use of a high-quality HF filter in combination with the switched capacitor subsampling topology is therefore absolutely necessary.

In Ref. [3], a fully integrated quadrature down-converter is presented. The circuit requires no external components, nor does it require tuning or trimming. It uses a double-quadrature structure, which renders a very high performance in quadrature accuracy. The down-converter topology is based on the use of MOS transistors in the linear region. By creating a virtual ground, a low frequency op-amp can be used for down conversion. The MOS transistor in the linear region results in a very high linearity for both the RF and the LO signal.



(a) Second-order intermodulation point



(b) Third-order intermodulation point

FIGURE 4.11 Linearity as a function of the gate overdrive voltage.

4.5 Synthesizer

One fundamental building block in every RF transceiver is the frequency synthesizer. The frequency synthesizer is responsible for generating the LO signal. The signal generated by the frequency synthesizer needs to be clean since low oscillator noise is crucial for the quality and reliability of the information

transfer. The signal should also be programmable and fast switching to be able to address all frequency channel within the specified time frame.

4.5.1 Topology

Synthesizers can usually be divided into three categories: table look-up synthesizer, the direct synthesizer, and the indirect or PLL synthesizer. In a table look-up synthesizer, the required sinusoidal frequency is created piece by piece using digital representations stored in memory of the amplitude at different time points of the sinusoidal waveform. The required building blocks are an accumulator that keeps track of the time, a memory containing a sine, a digital-to-analog converter (DAC), and a low-pass-filter to perform interpolation of the waveform to remove high frequency spurs. This type of synthesis is limited in frequency due to the access time of the memory and due to the maximum operation frequency of the high accuracy DAC. Moreover, high frequency spurs, generated due to the sampling behavior of the system, tend to corrupt the spectral purity of the signal. The direct frequency synthesizer employs multiplication, division, and mixing to generate the wanted frequency from a single reference. By repeatedly mixing and dividing, any accuracy is possible. The output spectrum is as clean as the reference frequency spectrum. Very fast frequency hopping is possible. The main disadvantages of this type of system is the difficult layout of the system, the high power consumption due to the numerous components, and the spectral purity can be corrupted by cross-coupling between stages. For generating high frequencies, the indirect or PLL type of frequency synthesizer often is the best choice. In a PLL, the synthesized frequency is generated by locking a locally generated frequency to an external frequency. The external frequency originates from a low frequency high quality crystal oscillator. To generate to local signal in the PLL, a VCO is used. A simple PLL topology is shown in Figure 4.12. A PLL includes following the building blocks: a VCO, a phase/frequency detector (PD/PFD), a loop filter, and a frequency divider or prescaler. The last building block is needed to derive a low frequency signal from the LO. This allows the signal to be locked to the external frequency through means of the PD. The PD is a circuit that compares the external frequency phase with the locally generated frequency phase and outputs an error voltage proportional to the phase difference. After filtering, this error signal is fed back to the VCO. This constitutes a control system. Under lock conditions, the external frequency and the locally generated frequency have a constant phase relationship.

$$F_{\text{out}} = N \cdot F_{\text{ref}} \tag{4.24}$$

The two signals are locked to each other, hence the name PLL. Even when a low-quality LO signal is generated, a high-quality signal can be synthesized. Due to the phase relationship between the input and the output frequency, the output signal will have the same spectral purity as the input high-quality signal. This is due to the fact that the loop remains locked to the input phase and therefore follows the phase deviations of that signal thus taking over its phase noise. This, however, is only true as long as the loop dynamics can follow the input signal. The loop dynamics are mainly determined by the bandwidth of the loop. For offset frequencies below the loop bandwidth, the phase noise is determined by the phase noise of the reference signal, for frequency offsets above the loop bandwidth, the output phase noise will be determined by the phase noise of the locally generated signal.

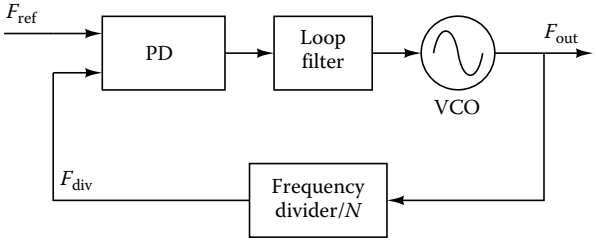


FIGURE 4.12 PLL-based frequency synthesizer.

When a programmable frequency divider is used in the loop, one can see that a set of frequencies can be synthesized. Suppose that the frequency by which the output signal is divided can be varied between N_1 and N_2, the output becomes

$$F_{out} = N_1 \cdot F_{ref},\ (N_1 + 1) \cdot F_{ref}, \ldots, N_2 \cdot F_{ref} \tag{4.25}$$

The PLL synthesizer is inherently slower than the other two types of synthesizers. The switching speed between two frequencies in Equation 4.25 is mainly determined by the loop bandwidth. Fast switching is only possible if a high loop bandwidth is implemented. Note that the loop bandwidth will also determine the phase noise performance. One, however, cannot indefinitely enlarge the loop bandwidth for stability reasons. A rule of thumb is that the loop bandwidth may not exceed 10% of the reference frequency to maintain stability. The loop bandwidth will also be limited by phase noise constraints. Spurious suppression and in-band phase noise levels will ultimately determine the loop bandwidth. When a low bandwidth has to be implemented, large capacitors will be needed. The total capacitance value is mainly determined by the need for implementing a stabilizing low frequency zero in the loop filter. This makes integration difficult as it will blow up silicon area and therefore increases the cost. One must therefore find ways to implement small bandwidth without having to use large capacitors. One obvious way of doing this is creating a low frequency pole through the use of a large resistance. This, however, will increase the phase noise. Other techniques, however, exist. In Ref. [35], a dual path loop filter is used. The filter consists of one active path and one passive path. Combining both will create a low frequency zero without the need for an extra resistor and capacitor. In Ref. [36], another technique is used to create the low frequency zero. It is created in the digital domain. The signal in the loop filter is combined with a sampled delayed version of itself. If the required switching speed is not achieved with a PLL configuration, one can make a combination of the direct synthesizer with the indirect synthesizer. In this topology, a number of PLLs are implemented and the outputs of all are combined with mixers. In this way, it is possible to synthesize a wide frequency range with a fast switching speed. This technique has recently been adopted for use in ultrawide band systems [37]. The major drawback of this technique, however, is that single sideband mixers have to be used. This requires accurate quadrature phases in all PLLs, low harmonic distortion, and well-matched mixers.

4.5.2 Oscillator

As it was mentioned above, the VCO is the main source of the phase noise outside the loop bandwidth. Therefore, its design is one of the critical parts of a PLL design. For the design of sub-gigahertz VCO, two oscillator types are often used: ringoscillators and oscillators based on a resonant tank composed of an inductor and a capacitor. The last type is referred to as an LC-tank VCO. The inductor in an LC-tank VCO can be implemented in two ways: an active implementation and a passive implementation. It can be shown [38,39] that the phase noise is inversely proportional to the power consumption. In LC-tank VCOs, the power consumption is proportional to the quality factor of the tank. Equations 4.26 through 4.28 show this relationship.

$$\text{Ring Osc. [39]}: \quad \mathcal{L}\{\Delta\omega\} \sim kTR \cdot \left(\frac{\omega}{\Delta\omega}\right)^2 \quad \text{with} \quad g_m = \frac{1}{R} \tag{4.26}$$

$$\text{Active LC [38]}: \quad \mathcal{L}\{\Delta\omega\} \sim \frac{kT}{2\omega C} \cdot \left(\frac{\omega}{\Delta\omega}\right)^2 \quad \text{with} \quad g_m = 2\omega C \tag{4.27}$$

$$\text{Passive LC [38]}: \quad \mathcal{L}\{\Delta\omega\} \sim kTR \cdot \left(\frac{\omega}{\Delta\omega}\right)^2 \quad \text{with} \quad g_m = R(\omega C)^2 \tag{4.28}$$

It is clear that for high frequency, a low power solution is only viable with an LC-tank VCO with a passive inductor. The use of a passive inductor, however, comes at a severe area penalty. Moreover, as it was discussed in Section 4.3, high-quality integrated inductors are difficult to make. For extremely low phase noise VCOs, bond wire inductors have been investigated [38]. The main drawback of using bondwires as inductors lies in reliability and yield. It is very difficult to make two bondwires exactly the same and reproduce this several times.

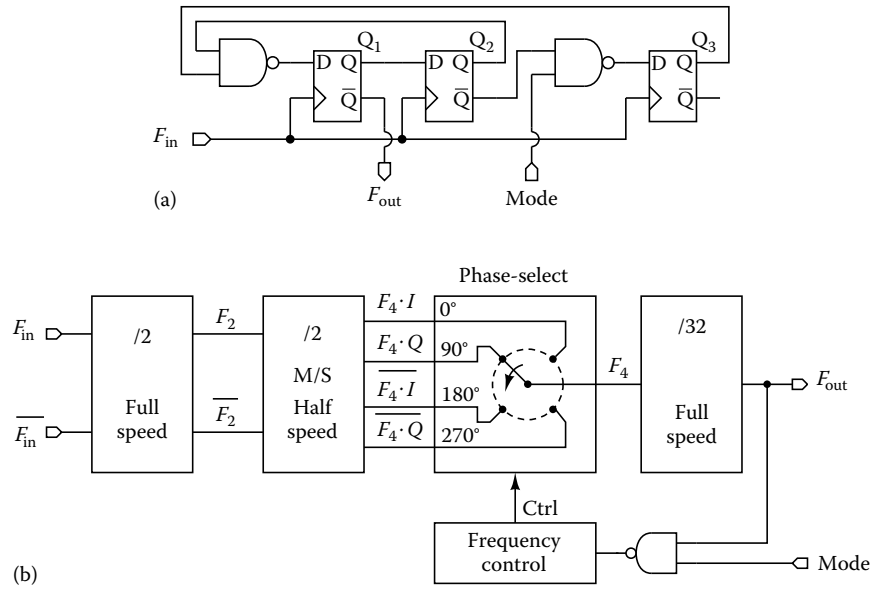


FIGURE 4.13 Dual modulus prescaler architecture: (a) D-flipflop based and (b) phase select topology. (From Craninckx, J. and Steyaert, M. *IEEE J. Solid-State Circuits*, 30(12), 1474, 1995.)

4.5.3 Prescaler

Several structures can be used as programmable divider. Programmable counters are the easiest solutions and are available in standard cell libraries. They are, however, limited in operation frequency. When high frequencies need to be synthesized, one can use a so-called prescaler. A prescaler divides by a fixed ratio and can therefore operate at high frequencies because they do not have to allow for delays involved with counting and presetting. A few high speed prescaller stages lower the speed used in the following counter stages. The disadvantage is that for a certain frequency resolution, the reference frequency has to be lowered. This slows the loop down as a lower bandwidth has to be implemented to maintain stability in the loop. A solution to this resolution problem is the use of dual- or multi-modulus prescalers. This circuit extends the prescaler with some extra logic to allow the prescaler to divide by N and $N+1$ in case of a dual-modulus prescaler and by N to $N+x$ in case of a multi-modulus prescaler. The speed decrease of this extra circuitry can usually be kept limited. Figure 4.13 shows two possible implementations of a dual-modulus prescaler. Implementation given by Figure 4.13a is a straightforward implementation based on d-flipflops. The critical path consists of a NAND gate and a d-flipflop. Implementation given by Figure 4.13b is a more complex implementation. It is based on the 90° phase relationship between the outputs of a master/slave toggle flipflop. It contains no additional logic in the high frequency path. The dual-modulus prescaler is as fast as an asynchronous fixed divider.

4.5.4 Fractional-*N* Synthesis

As it can be concluded from Equation 4.25, the minimal frequency resolution that can be achieved when using the topologies described previously is equal to F_{ref}. In GSM, e.g., the channels are 200 kHz spaced apart, this means that we need a frequency resolution of 200 kHz and therefore a low reference frequency. This results in high division ratios. The in-band phase noise of a PLL, however, is proportional to the division ratio, large ratios mean high in-band noise. As it is already explained, a low reference frequency will also result in a low PLL bandwidth and therefore a slow loop. Therefore, we need a technique that enables us to use a high reference frequency and still achieve the required frequency resolution. Fractional-N synthesizers solve this problem. Figure 4.14 makes things clearer. A basic fractional-N

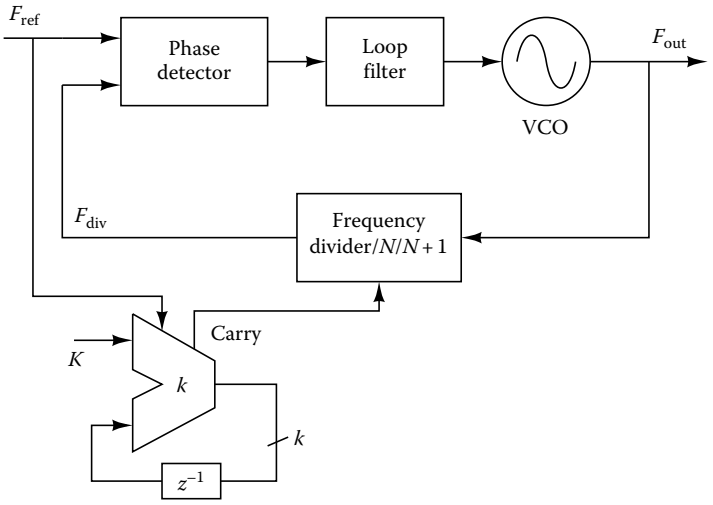


FIGURE 4.14 Fractional-N principle.

synthesizer consists, besides the standard PLL building blocks, of an accumulator and a dual modulus prescaler. By switching fast between the two division ratios, fractional divisions can be synthesized. The accumulator increases its value every reference clock cycle with a certain amount $K = n \cdot 2^k$. The dual-modulus prescaler is controlled by the accumulator overflow bit. If the accumulator overflows, the division ratio is $N+1$, otherwise it is N. On average, the dual-modulus prescaler divides K times by $N+1$ and $2^k - K$ times by N, resulting in a synthesized frequency of

$$N_{\text{frac}} = \frac{(2^k - K) \cdot N + K \cdot (N+1)}{2^k} = N + \frac{K}{2^k} = N + n \tag{4.29}$$

This means that also non-integer ratios can be synthesized and the above-mentioned limitations on the reference frequency is not applicable. There are, of course, drawbacks to the technique. The major one is the generation of spurs in the output spectrum due to pattern noise in the overflow signal. A detailed study of fractional-N synthesis, however, is beyond the scope of this chapter and the reader is referred to the open literature for further information. A thorough study of fractional-N synthesizers and their simulation can be found in Ref. [41].

4.6 Transmitter

Most RF communication systems are based on bidirectional data traffic. This means that apart from the receiver section, also a transmitter section must be implemented to complete a full transceiver. As explained in Section 4.2, a transmitter commonly includes a number of mixers, LO, and a PA. The LO is covered in a previous section; this section will therefore only describe the mixer and the PA used in up-conversion or transmitter systems.

4.6.1 Up versus Down Conversion

Although a huge amount of literature exists concerning the down-conversion process, up conversion has long time been neglected. This is rather surprising. When looking in Figure 4.5, one immediately notices

the parallelism between the receiver and the transmitter. The same functionality occurs. Both paths contain an amplifier (LNA, PA), both contain an interface to the digital domain (A/D, D/A), both contain a mixer and both are steered by the same LO system. The nature of the signals in both paths (input and output) has a huge influence on the circuit implementation. This seems logical for the LNA/PA analogy or the A/D and D/A-converter. Both have completely different topologies. Although the mixers in the up-conversion path and the down-conversion path face the same signals, there is typically not a great difference between the up- and down-conversion mixer topology. Most implementations are variations on the four-quadrant mixer topology, better known as the Gilbert-cell [42]. There are, however, fundamental differences between up conversion and down conversion. The first fundamental difference is located in the input signals of the mixer. In case of a down-conversion mixer, the input usually is a high-frequency, low-power signal surrounded by large blocking signals. In case of an up conversion, the input signal is a locally generated large baseband signal with a clean spectrum. At the output side, the situation is the opposite. A down-converted signal is a low frequency signal. It is, therefore, relatively easy to filter or apply feedback to cope with unwanted signals. At the transmitter side, however, a large and linear signal has to be processed within the technology dependend limited frequency range. Every extra building block placed between the mixer and the PA has to deal with high-frequency signals. Filtering is, therefore, impossible behind the up-conversion mixer as it will require large amount of power. Therefore LO leakage and other unwanted signals like intermodulation products have to be limited. A last, but not least difference lies within one of the design specifications of a mixer, the conversion gain G_c. It is defined as the ratio between the input power of the mixer and the output power. At the receiver side, the mixer input power is a design constraint as it is determined by the application. At the transmitter side, both input and output power are design variables. They can both be chosen freely. As it is easier and more power friendly to amplify a low-frequency signal, a large baseband signal is preferred.

4.6.2 CMOS Mixer Topologies

4.6.2.1 Switching Modulators

Many mixer implementations are based on the traditional variable transconductance multiplier with cross-coupled differential modulator stages [42]. It is depicted in Figure 4.15. The circuit was originally implemented in a bipolar technology and therefore based on its inherent translinear behavior. The MOS counterpart, however, can only be effectively used in switching mode. This induces the use of large LO driving signals and result in large LO feedthrough and power consumption. Moreover, when using a square-wave type modulation signal, a lot of energy is located at the third harmonic.

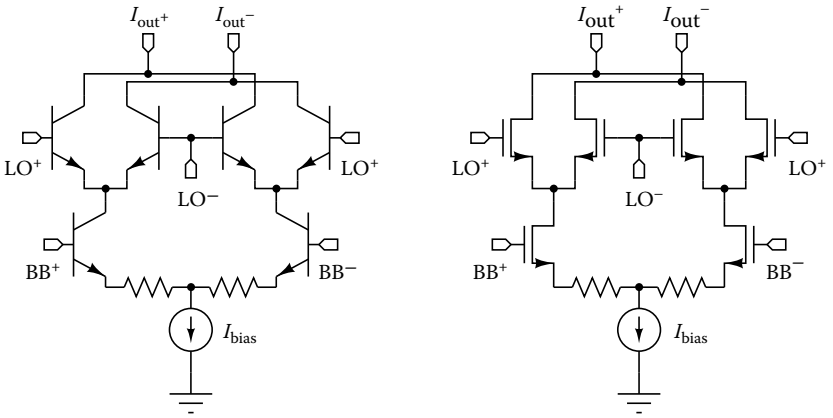


FIGURE 4.15 Bipolar and CMOS version of the Gilbert cell mixer.

This unwanted signal can only be filtered out by an extra blocking filter at the output. In CMOS the variable transconductance is typically implemented using a differential pair biased in the saturation region. To avoid distortion problems, large $V_{gs} - V_T$ values or a large source degeneration resistor is needed. This results in a large power consumption and noise problems. For upconversion, one also has to be aware that the high frequency current has to run through the modulating transistors. The source degeneration is therefore limited by bandwidth constraints. These problems can be circumvented by replacing the bottom differential pair with a pseudo-differential topology biased in the linear region [43].

4.6.2.2 Linear MOS Mixers

Figure 4.16 presents a linear CMOS mixer topology together with an output driver [44,45]. The circuit implements a real single-ended output topology avoiding the use of external combining. Some basic design ideas and some guidelines to optimize the circuit will be presented. The circuit is based on an intrinsically linear mixer topology. The circuit feature four mixer transistors biased in the linear region. Each mixer converts a quadrature LO voltage and a baseband signal to a linearly modulated current. The expression for the source-drain current for an MOS transistor in the linear region is given by

$$I_{DS} = \beta\left[(V_{GS} - V_T)V_{DS} - \frac{V_{DS}^2}{2}\right] \tag{4.30}$$

This equation can be rewritten in terms of a DC and an AC term:

$$\begin{aligned} I_{DS} &= \beta(V_{DS} + v_{ds}) \cdot \left(V_{GS} - V_T - \frac{V_D - V_S}{2} + v_g - \frac{v_d + v_s}{2}\right) \\ &= \underbrace{\beta V_{DS}\left(V_{GS} - V_T - \frac{V_D - V_S}{2}\right)}_{\text{DC component}} \\ &\quad + \underbrace{\beta v_{ds}\left(V_{GS} - V_T - \frac{V_D - V_S}{2}\right) + \beta V_{DS}\left(v_g - \frac{v_d + v_s}{2}\right) + \beta v_{ds}\left(v_g - \frac{v_d + v_s}{2}\right)}_{\text{AC component}} \end{aligned} \tag{4.31}$$

Two signal have to be applied to a mixer transistor, the low frequency baseband signal, and the high frequency LO signal. Applying these signals may only result in the wanted high-frequency currents. Based on Equation 4.31, some conclusions can be drawn.

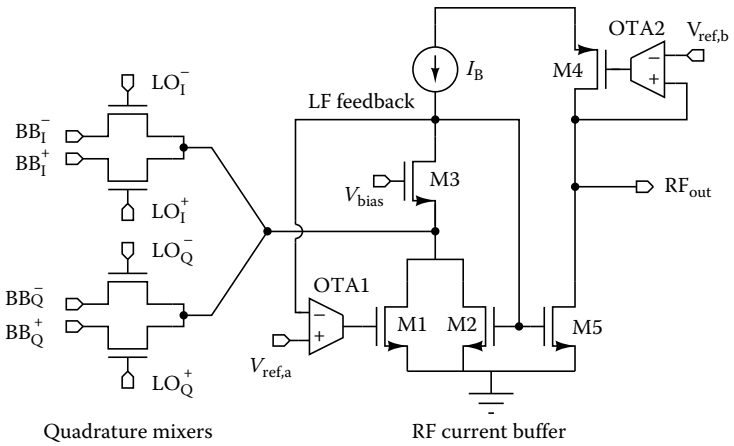


FIGURE 4.16 Schematic of a linear up-conversion mixer with output driver.

If the LO signal is applied to the drain/source of the mixer transistor, a product term

$$\beta v_{ds} \cdot \left(V_{GS} - V_T - \frac{V_D - V_S}{2} \right)$$

is formed. As this contains the product of a DC voltage with the oscillator signal, this component is located at LO frequency. It is preferable to avoid this frequency component to be formed. Therefore, the LO signal should not be applied to this node. Applying the LO signal to the gate of the mixer transistors results in the wanted behavior. According to Equation 4.31, only the high-frequency components are formed by

$$\beta v_g \cdot (V_{DS} + v_{ds})$$

By applying a zero-DC voltage between source and drain, only the high frequency mixer product is generated. The voltage to current conversion is perfectly balanced. The current of the four mixer paths is immediately added at the output of the mixers at a common node. This requires a virtual ground at that point that is achieved due to the low impedance input of the buffer stage (Figure 4.16). The total current flowing into the output buffer is given by

$$I_{MIX} = \beta \left(v_{bb_I}^2 + v_{bb_Q}^2 + 2 \cdot v_{LO_I} v_{bb_I} + 2 \cdot v_{LO_Q} v_{bb_Q} \right) \tag{4.32}$$

Equation 4.32 shows two frequency components in the modulated waveform. $\beta v_{LO} v_{bb}$ is the wanted signal. To prevent intermodulation products of the low frequency baseband signal $\beta v_{bb_I}^2$ with the wanted RF signal, the LF signal has to be suppressed at the current summing node. This is achieved by a low frequency feedback loop in the output buffer.

The low frequency feedback loop consists of OTA1 and transistors M1 and M3. It suppresses the low frequency signals resulting in a higher dynamic range of the output stage and decreases unwanted intermodulation products. It also lowers the input impedance of the output stage at low frequencies. The structure in fact separates the high- and low-frequency components of the input current and prevents the low frequency component to be mirrored to the output stage. The RF current buffer also ensures a low impedance at high frequencies at the mixer current summing node and therefore provides the necessary virtual ground.

4.6.2.3 Nonlinearity and LO Feedthrough Analysis

The difficulty to integrate IF filters is one of the reasons to implement direct conversion transmitters. This implies the the LO is at the same frequency as the RF signal and cannot be filtered out. To minimize the spurious signal components at the LO frequency, one has to isolate the origins of the unwanted frequency components. They can be categorized in three topics: capacitive feedthrough due to gate-source and gate-drain, parasitic overlap capacitances, and intrinsic nonlinearity of the mixers, mixer products due to a nonideal virtual ground.

When an ideal virtual ground is provided at the output of the mixer, capacitive LO feedthrough is canceled. This cancellation is never perfect, however, due to technology mismatch. The capacitive LO feedthrough for a single mixer transistor, biased in the linear region, is therefore given by

$$I_{LO} = 2\pi f \cdot v_{LO} \cdot WL \cdot \left(\frac{C_{ox}}{2} + \frac{C_{ov}}{L} \right) \tag{4.33}$$

where

c_{ox} is the oxide capacitance
C_{ov} is the gate-drain/source overlap capacitance
v_{LO} is the amplitude of the LO signal
f is its frequency

Based on Equations 4.32 and 4.33, the ratio between the LO feedthrough current and the modulated current is given by

$$\frac{i_{\text{signal}}}{\Delta(i_{\text{LO}})} = \frac{2\mu C_{\text{ox}} \frac{W}{L} v_{\text{bb}} v_{\text{LO}}}{\delta(i_{\text{LO}}) \cdot 2\pi f \cdot v_{\text{LO}} \cdot WL \cdot \left(\frac{C_{\text{ox}}}{2} + \frac{C_{\text{ov}}}{L}\right)} = \frac{\mu C_{\text{ox}} v_{\text{bb}}}{\delta(i_{\text{LO}}) \cdot \pi f \cdot L^2 \cdot \left(\frac{C_{\text{ox}}}{2} + \frac{C_{\text{ov}}}{L}\right)} \tag{4.34}$$

where $\delta(i_{\text{LO}})$ accounts for the relative difference in LO feedthrough for the different mixer transistors due to mismatch. Equation 4.34 shows that the ratio between the modulated current and the LO feedthrough current is independent from the LO amplitude and from the transistor width. Feedthrough will be less if shorter transistor lengths are used. The relative matching between the different mixer transistor will become worse, however, when shorter lengths are used [46]. One must therefore use Equation 4.34 with care. The δ will increase for smaller transistor length. With proper design and optimization, one should, however, be able to achieve a 30 dBC signal to LO feedthrough ratio even if two LO feedthrough currents are added instead of being canceled by the virtual ground ($\delta(i_{\text{LO}} = 1)$). When more realistic numbers of mismatch are considered (e.g., 10%), 50 dBC is easily achieved. The presented equations can, therefore, be used by the experienced designer to estimate the matching requirements and check if these requirements are realistic.

Another problem one faces is a possible DC-offset between the source and drain terminal of the mixer transistor. Equation 4.31 explains the problem. Ideally, no DC is present. The mixer then shows the required behavior. When a DC is present, however, one can see that components are generated at DC, the LO frequency due to multiplication with v_{g} and a component at baseband. While the low frequency components can be filtered out by the low frequency feedback in the output buffer, the component at the LO frequency remains. This component will, therefore, set the requirements for the tolerated DC-offset. A possible solution for this problem is measuring the DC-offset between source and drain. The offset is then controllable. The offset requirement is translated into an offset specification on the op-amps used in the feedback loops in the output buffer.

If the common mixer node is not a ideal virtual ground, the modulated current will be converted to a voltage dependent on the impedance seen on that node. The spectrum of the modulated signal will therefore be a combination of the modulated current spectrum and the frequency dependence of the impedance. When an impedance Z_{c} is considered at the common mixer node, the modulated current is given as the result of a second-order equation

$$\left(\beta Z_{\text{c}}^2\right) \cdot I^2 - \left(1 + 2\beta Z_{\text{c}} \cdot (V_{\text{GS}} - V_{\text{T}})\right) \cdot I + 2\beta v_1 v_{\text{g}} + \beta v_1^2 = 0 \tag{4.35}$$

It can be noticed that when $Z_{\text{c}} = 0$, Equation 4.35 is reduced to Equation 4.32. As Equation 4.35 is a second-order equation, it is a possible origin of distortion and therefore has to be taken into account. One side note should be made to the previous. Only currents that are not canceled out by the differential character of the mixer are converted in a voltage. This advantage of a balanced structure, however, is not valid for a nonideal voltage source at the input of the mixer transistors. If a nonideal voltage source is used at this node, each frequency component of the modulated current will be converted in a voltage according to the specific frequency-dependent impedance. These voltages are then similar to the baseband signal up-converted to the LO frequency. It is therefore essential to keep this node as low-impedance as possible.

Equation 4.32 is only valid if a very low impedance is seen at the source and drain terminals of the mixer transistors. If this condition is fulfilled, no unwanted high-frequency mixing components are present in the modulated signal. However, both in measurements and in simulations, a significant unwanted signal is noticed at $f_{\text{LO}} \pm 3f_{\text{bb}}$. One expects this component to originate from a $v_{\text{LO}} v_{\text{bb}}^3$ product term. However, Equation 4.35 only shows a second-order relationship. The observed product term must therefore find its origin in another effect. It is proved to be a result of short channel effects in an

MOS transistor. Both the effective mobility and the threshold voltage are affected by the gate-source and drain-source voltage. The calculated impact of the threshold voltage modulation cannot explain the observed effect; it is therefore assumed that it is a result of the mobility modulation. After some calculations, one can prove that the effective mobility is

$$\mu_{\text{eff}} = \frac{\mu_0}{1 + \theta \cdot (V_{\text{GS}} - V_{\text{T}})_{\text{dc}} + \theta \cdot \left(v_{\text{LO}} - \frac{v_{\text{bb}}}{2}\right) + \frac{\mu_0}{V_{\max} \cdot L} + \frac{\theta}{2} \cdot |v_{\text{bb}}|} \tag{4.36}$$

Substituting v_{bb} with $A\sin(\omega_{\text{bb}}t)$ and making a Fourier series expansion of $|v_{\text{bb}}|$ results in

$$\mu_{\text{eff}} = \frac{\mu_0}{B\left(1 + \frac{\theta}{B} v_{\text{LO}} - \frac{\theta A}{2B} \sin(\omega_{\text{bb}}t) + C\cos(2\omega_{\text{bb}}t) + D\cos(4\omega_{\text{bb}}t) + \cdots\right)} \tag{4.37}$$

with

$$\begin{aligned} A\sin(\omega_{\text{bb}}t) &= \text{the baseband signal} \\ B &\approx 1 + (V_{\text{GS}} - V_{\text{T}}) \\ C &= \frac{A}{B} \cdot \frac{4}{3\pi} \cdot \left(\frac{\mu_0}{V_{\max} \cdot L} + \frac{\theta}{2}\right) \\ D &= \frac{A}{B} \cdot \frac{4}{5 \cdot 3\pi} \cdot \left(\frac{\mu_0}{V_{\max} \cdot L} + \frac{\theta}{2}\right) \end{aligned}$$

Equation 4.37 shows that a second-order baseband frequency component $\cos(2\omega_{\text{bb}}t)$ appears. In the DC reduction factor B, the third term is an order of magnitude smaller than 1. Hence it appears that the magnitude C of the second-order component has a first order relationship to the baseband signal amplitude A. In the voltage to current relationship, μ_{eff} is multiplied with $v_{\text{LO}}v_{\text{bb}}$. As a result a mixing component at $f_{\text{LO}} \pm 3f_{\text{bb}}$ occurs. In the amplitude C of this distortion component, $\mu_0/(V_{\max} \cdot L)$ is dominant to $\theta/2$ for most submicron technologies. It is also important to notice that the distortion is inversely proportional to the gate length. This indicates that this effect will become even more apparent as gate lengths continue to scale down.

4.6.3 Power Amplifier

4.6.3.1 CMOS Power Amplification

The integration of PAs in a CMOS technology is impeded by the low supply voltage of the current deep-submicron and nanometer technologies. Apart from this, the relative high parasitic capacitances of the MOS transistor, at least compared to GaAs or SiGe transistors, and the relative low quality factor of on-chip inductors, further hinders the integration. On the other hand, the digital MOS transistor is optimized for switching and as such a lot of switching amplifiers have been integrated in CMOS with great success recently [47–53]. Furthermore, CMOS RF amplifiers are capable to break the 1 W barrier of output power performance [54]. In this section, the topic of switching RF amplifier is discussed first. In the second part, some linearization techniques will be discussed.

4.6.3.2 Switching Class E Amplifier

The Class E amplifier was invented in 1975 [55], but the first implementation of this amplifier in CMOS was reported in 1997 [47]. In contrast to the Class A, B, C, and F amplifiers, the Class E is designed in the time-domain. In theory, the Class E amplifier is capable to achieve an efficiency of 100%. In order to achieve this, the transistor and output network are designed in such a way that the drain through the transistor is separated in time from the voltage across the transistor. This avoids power dissipation in the transistor, a necessary requirement to achieve a high efficiency. If all other elements are assumed to be

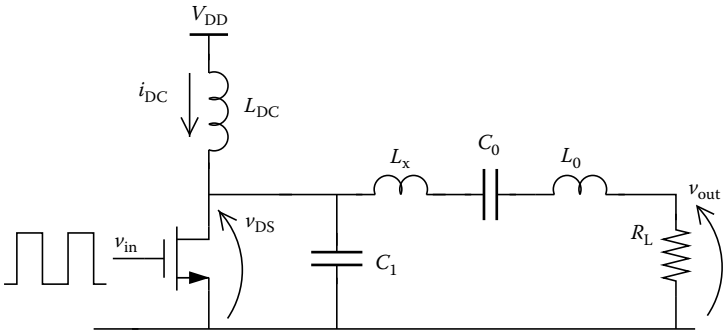


FIGURE 4.17 Basic Class E PA.

lossless, the amplifier is then indeed capable to achieve an efficiency of 100%. Figure 4.17 depicts the basic circuit of a CMOS Class E amplifier. The nMOS transistor should act as a switch and therefore it is driven by a square wave between zero and the maximum permissible gate voltage, which is normally equal to V_{DD}, the supply voltage of the technology. As such the nMOS transistor can be modeled by an ideal switch with a series resistance R_{on}. Inductor L_1 can be seen as the DC feed inductance, and in the original Class E theory, this inductor is assumed to be very large, and can be replaced by an ideal current source. Finally, inductor L_x and capacitor C_1 are the two crucial elements that create the Class E waveform at the drain of the nMOS transistor.

In a fully integrated CMOS implementation, the DC feed inductor L_1 cannot be made very large. First, this would require a huge silicon area, but more important, the relative high power loss of CMOS integrated inductors does not allow for such a large value. As such, the value of L_1 has to be reduced, and the current through the latter will not be a constant. The amplifier can still be designed to meet the Class E conditions, even with a small value of L_1. In fact, reducing the value of L_1 will result in a larger value for C_1 and a smaller value for L_x.

The capacitor C_1 and inductor L_x are constrained by the two Class E requirements, given below.

$$\text{Class E} \Leftrightarrow \begin{cases} v_{DS}(t = t_1) & = & 0 \\ \left.\dfrac{dv_{DS}(t)}{dt}\right|_{t=t_1} & = & 0 \end{cases}$$

In Figure 4.18, the drain voltage and current for Class E operation are shown. Solving the two Class E equations will give a value for C_1 and L_x. Finally, the value of the load resistance is constrained by the required output power. To achieve sufficient output power in a low voltage CMOS technology, an impedance matching network is required between the 50 Ω load or antenna impedance and the Class E amplifier. The on resistance of the nMOS transistor can be written as

$$r_{on} = \frac{L}{\mu_n C_{ox} W (V_{GS} - V_T)} \tag{4.38}$$

The lower the on resistance, the higher the efficiency of the amplifier, and thus it is beneficial to increase the width of the nMOS transistor. However, that large transistor cannot be directly connected to the up-conversion mixer, and several amplifying stages are needed between them. If the nMOS transistor has a large gate width, more power will be consumed by the driver stages and thus the overall efficiency of the amplifier, defined as

$$\eta_{oa} = \frac{P_{out}}{P_{DC,PA} + P_{DC,DRV}} \tag{4.39}$$

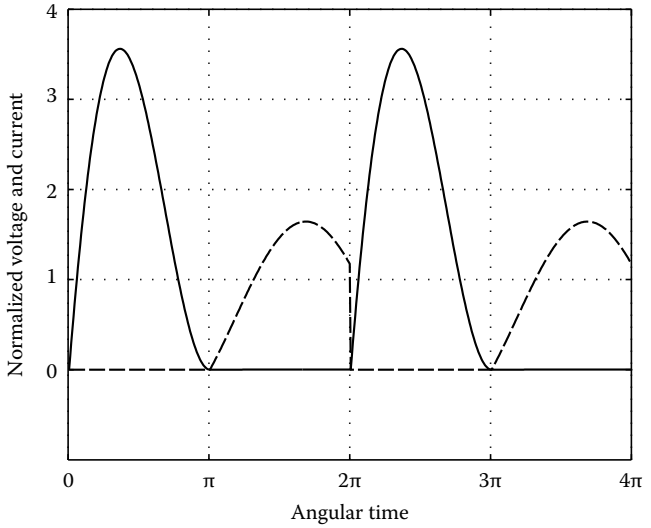


FIGURE 4.18 Voltage (solid line) and current (dashed line) of the Class E PA.

will have a maximum value for a specific transistor width. The overall efficiency is not always a good figure to compare PAs, since that figure can never reach 100%, even if each of the stages has a conversion efficiency of 100%. After all, the power consumed by the driver stages will never flow to the output load, but is only needed to switch on and of the next stage in line.

The power added efficiency (PAE) defined as

$$\text{PAE} = \frac{P_{\text{out}} - P_{\text{in}}}{P_{\text{DC}}} \tag{4.40}$$

is a useful definition for stand-alone PAs that have an input matched to 50 Ω. However, one should be aware whether the DC power consumption of the driver stages is included in P_{DC}.

Another important aspect of switching amplifiers is the reliability. A drawback of the Class E amplifier, at least compared to the Classes B and F, is that the drain voltage goes up to several times the supply voltage of the amplifier. This might cause reliability problems. On the other hand, the switching nature of the amplifier alleviates this. After all, due to the switching, voltage and current are separated in time. In other words, the high voltage peaks are not accompanied by a drain current, and when the drain current is high, the voltage across the switch is low. This is a big advantage compared to other types of amplifiers.

Figure 4.17 depicts another benefit of the Class E amplifier. For Class E operation, a shunt capacitance C_1 is required at the drain. However, in CMOS, there is already a large parasitic drain capacitance, and this capacitance can now become part of the amplifier. In Classes B and F amplifiers, that parasitic capacitance will create a low impedance for the harmonics that are crucial for the high efficiency of Classes B and F. Therefore, CMOS seems to be the natural habitat of the Class E amplifier.

4.6.3.3 Linearization of CMOS RF PAs

Switching amplifiers only have phase linearity, and therefore are only useful for constant envelope systems like Bluetooth and GSM. However, modern RF communication systems like UMTS, CDMA-2000, and WLAN allow amplitude modulation to increase the datarate of a wireless link. The only way to recover or restore the amplitude linearity of a switching amplifier is by modulating the supply voltage or by combining two nonlinear amplifiers. Systems that modulate the supply voltage of a switching

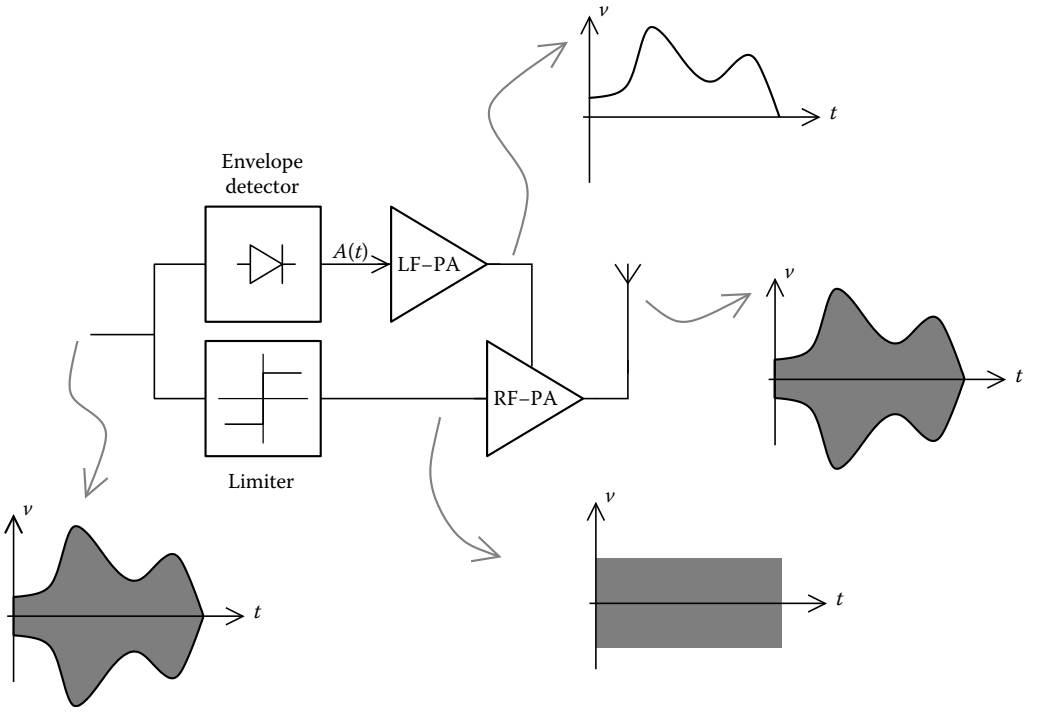


FIGURE 4.19 Kahn technique linearized PA.

amplifier are denoted as ''envelope elimination and restoration,'' ''polar modulation,'' or ''supply modulation.'' They originate from the Kahn technique (see Figure 4.19) that was already employed in vacuum tube amplifiers. In CMOS, one can make use of the availability of digital signal processing to directly create the amplitude and phase signal, and as such, the limiter and envelope detector of Figure 4.19 can be avoided. Furthermore, AM-AM and AM-PM predistortion is relative easy to implement. The general picture of polar modulation is shown in Figure 4.20. An another advantage of polar-modulated amplifiers is that the entire phase path carries a constant envelope signal and thus one can use nonlinear or saturated blocks in the upconversion path. Furthermore, amplitude and phase feedback are relatively easy to implement. Another group of techniques combine the output of two constant envelope amplifiers that have a different in phase. The two amplifiers are combined through a transformer, a power combiner, or through transmission lines, and the output is, in general, the sum of the two amplifiers. These systems are called ''outphasing'' or ''LINC,'' depending on the used combiner. Depending on their phase difference, the resulting output envelope can be higher or lower, and thus has amplitude modulation, as shown in Figure 4.21. The major drawback of these techniques is the difficulty to implement the power combiner in CMOS. Also, feedback is not as easy to implement in these systems. On the other hand, these systems allow to efficiently amplify signals that have a very high modulation bandwidth.

Apart from the two groups discussed in this section, several other techniques exist to amplify an amplitude-modulated signal. There is no ''ideal'' solution for CMOS integration. An alternative solution or approach is to use a linear amplifier with an efficiency improvement technique, such as the Doherty amplifier or the bias adoption technique. However, the linearization of nonlinear amplifiers has the advantage that switching or nonlinear amplifiers can be used, which are easier to implement in CMOS. Furthermore, the RF driver stages and all the blocks preceding the RF amplifier can be nonlinear as well. Needless to say, this is a huge advantage in low voltage technologies.

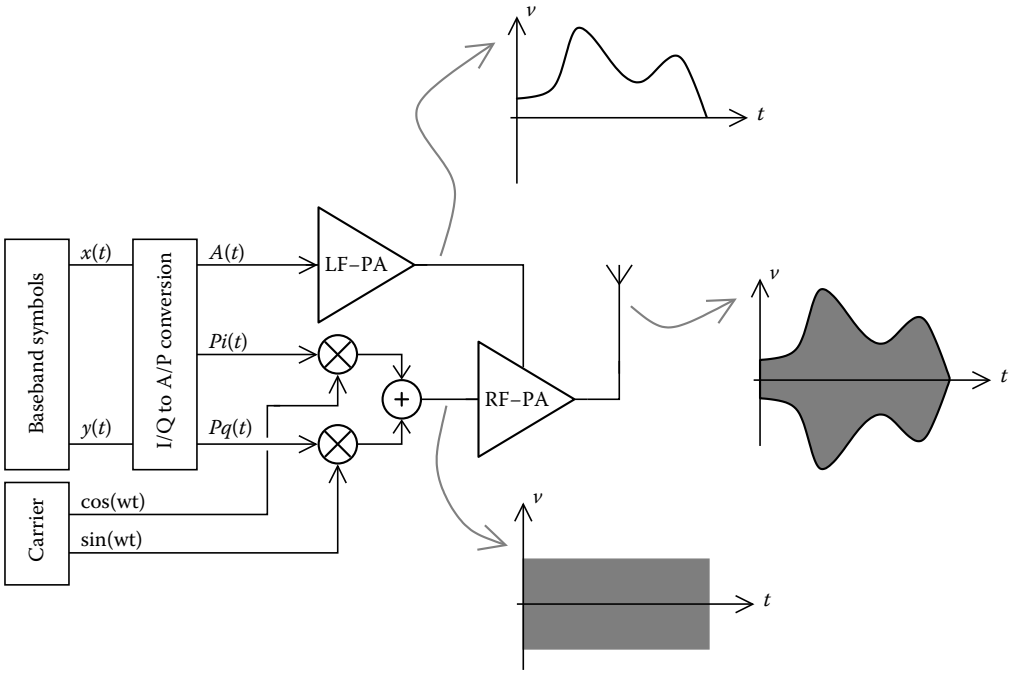


FIGURE 4.20 DSP based polar modulation architecture.

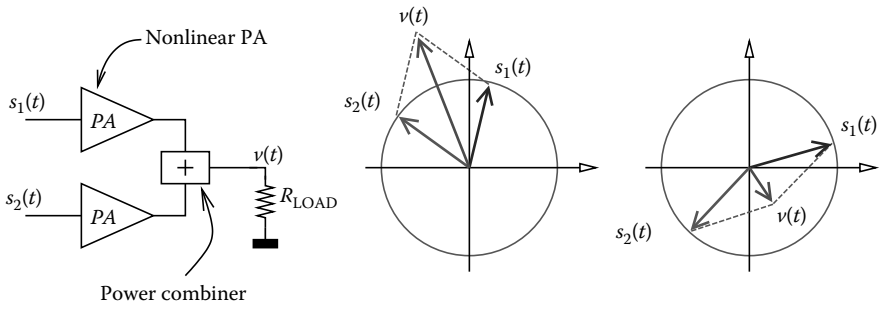


FIGURE 4.21 LINC or outphasing architecture.

References

1. C. E. Shannon, Communication in the presence of noise, in *Proceedings of the IRE*, vol. 37, Jan. 1949, pp. 10–21.
2. J. Sevenhans, A. Vanwelsenaers, J. Wenin, and J. Baro, An integrated Si bipolar transceiver for a zero-IF 900MHZ GSM digital mobile radio front-end of a hand portable phone, in *Proceedings of the Custom Integrated Circuits Conference, CICC*, May 1991, pp. 7.7.1–7.7.4.
3. J. Crols and M. Steyaert, A single Chip 900MHz CMOS receiver front-end with a high performance low-IF topology, *IEEE J. Solid-State Circuits*, 30(12), 1483–1492, 1995.
4. P. R. Gray and R. G. Meyer, Future directions in silicon ICs for RF personal communications, in *Proceedings of the Custom Integrated Circuits Conference, CICC*, May 1995, pp. 91–95.

5. A. A. Abidi, Low-power radio-frequency ICs for portable communications, *Proc. IEEE*, 83(4), 544–569, 1995.
6. S. Mehta, D. Weber, M. Terrovitis, K. Onodera, M. Mack, B. Kaczynski, H. Samacati, S. Jen, W. Si, M. Lee, K. Singh, S. Mendis, P. Husted, N. Zhang, B. McFarland, D. Su, T. Meng, and B. Wooley, An 802.11g WLAN SoC, in *ISSCC Digest of Technical Papers*, vol. 48, Feb. 2005, pp. 94–95.
7. H. Darabi, S. Khorram, Z. Zhou, T. Li, and B. Marholev, A fully integrated SoC for 802.11b in 0.18μm CMOS, in *ISSCC Digest of Technical Papers*, vol. 48, Feb. 2005, pp. 96–97.
8. National Telecommunications and Information Administration, Office of Spectrum Management (NTIA-OSM), http://www.ntia.doc.gov/osmhome/osmhome.html.
9. Federal Communications Commission (FCC), http://www.fcc.gov/.
10. R. L. Freeman, *Radio System Design for Telecommunications*, 2nd ed., Wiley, New York, 1997.
11. T. S. Rappaport, *Wireless Communications: Principles and Practice (Communication, Engineering, and Emerging Technologies)*, 2nd ed., Prentice Hall, Upper Saddle River, NJ, 2001.
12. L. W. Couch, *Digital and Analog Communication System*, Prentice Hall, Upper Saddle River, NJ, 1997.
13. A. Abidi, Direct conversion radio transceivers for digital communications, *IEEE J. Solid-State Circuits*, 30(12), 1399–1410, 1995.
14. B. Razavi, Design considerations for direct conversion receivers, *IEEE Trans. Circuits Syst. II*, 44(6), 428–453, 1997.
15. J. Janssens and M. Steyaert, *CMOS Cellular Receiver Front-Ends: From Specification to Realization (The International Series in Engineering and Computer Science)* Springer, The Netherlands, 2002.
16. K. Kurokawa, Injection locking of microwave solid-state oscillators, *Proc. IEEE*, 61, 1386–1410, Oct. 1973.
17. B. Razavi, A study of injection locking and pulling in oscillators, *IEEE J. Solid-State Circuits*, 39(9), 1415–1424, 2004.
18. M. Feulner, Direct up-conversion lowers base-station costs, *Wireless Europe Magazine*, April–May 2005, pp. 22–25, wireless.iop.org.
19. W. Liu, X. Jin, X. Xi, J. Chen, M.-C. Jeng, Z. Liu, Y. Cheng, K. Chen, M. Chan, K. Hui, J. Huang, R. Tu, P. K. Ko, and C. Hu, *BSIM3v3.3 MOSFET Model, User's Manual*, University of California, Berkeley, 1999.
20. M. Steyaert and W. Sansen, Opamp design towards maximum gain-bandwidth, in *Proceedings of the AACD Workshop*, Delft, Mar. 1993, pp. 63–85.
21. C. Enz and Y. Cheng, MOS transistor modeling for RF IC design, *IEEE J. Solid-State Circuits*, 35(2), 186–201, 2000.
22. J. Janssens, J. Crols, and M. Steyaert, Design of broadband low-noise amplifiers in deep submicron CMOS technology, in *Analog Circuit Design. 1 Volt Electronics, Mixed-Mode Systems, Low-Noise and RF Power Amplifiers for Telecommunication*, J. Huijsing, R. Van de Plassche, and W. Sansen, Eds., Kluwer Academic Publishers, Amsterdam, the Netherlands, 1999, pp. 317–335.
23. Y. P. Tsividis, *Operation and Modelling of the MOS Transistor*, McGraw-Hill, New York, 1987.
24. D. Leenaerts, J. van der Tang, and C. Vaucher, *Circuit Design for RF Transceivers*, Kluwer Academic Publishers, Boston, MA, 2001.
25. J. Craninckx and M. Steyaert, A 1.8GHz low phase noise CMOS VCO using optimized hollow spiral inductors, *IEEE J. Solid-State Circuits*, 32(5), 736–744, 1997.
26. N. Itoh, B. D. Muer, and M. Steyaert, Low supply voltage integrated CMOS VCO with three terminals spiral inductor, in *Proceedings of the European Solid State Circuits Conference, ESSCIRC*, Sept. 1999, pp. 194–197.
27. A.-S. Porret, T. Melly, C. C. Enz, and E. A. Vittoz, Design of high-Q varactors for low-power wireless applications using a standard CMOS process, *IEEE J. Solid-State Circuits*, 35(3), 337–345, 2000.
28. O. Kenneth, Estimation methods for quality factors of inductors fabricated in silicon integrated circuit process technologies, *IEEE J. Solid-State Circuits*, 33(8), 1249–1252, 1998.

29. C. S. Meyer, D. K. Lynn, and D. J. Hamilton, *Analysis and Design of Integrated Circuits*. McGraw-Hill, New York, 1968.
30. J. Crols, P. Kinget, J. Craninckx, and M. Steyaert, An analytical model of planar inductors on lowly doped silicon substrates for high frequency analog design to 3 GHz, in *Digest of Technical Papers, Symposium on VLSI Circuits*, 1996.
31. P. Leroux and M. Steyaert, *LNA-ESD Co-Design for Fully Integrated CMOS Wireless Receivers*, The International Series in Engineering and Computer Science, vol. 843, Springer, The Netherlands, 2005.
32. D. Rabaey and J. Sevenhans, The challenges for analog circuit design in mobile radio VLSI chips, in *Proceedings of the AACD Workshop* (*Leuven*), Mar. 1993, pp. 225–236.
33. D. Shen, H. Chien-Meen, B. Lusignan, and B. Wooley, A 900 MHz integrated discrete-time filtering RF front-end, in *ISSCC Digest of Technical Papers*, San Francisco, Feb. 1996, pp. 54–55, 417.
34. S. Sheng, L. Lynn, J. Peroulas, K. Stone, I. O'Donnell, and R. Brodersen, A low-power CMOS chipset for spread spectrum communications, in *ISSCC Digest of Technical Papers*, San Francisco, Feb. 1996, pp. 346–347, 471.
35. J. Craninckx and M. Steyaert, A fully integrated CMOS DCS-1800 frequency synthesizer, in *ISSCC Digest of Technical Papers*, San Francisco, Feb. 1998, pp. 372–373.
36. B. Zhang, P. Allen, and J. Huard, A fast switching PLL frequency synthesizer with an on-chip passive discrete-time loop filter in 0.25 μm CMOS, *IEEE J. Solid-State Circuits*, 38(6), 855–865, 2003.
37. D. Leenaerts1, R. van de Beek1, G. van der Weide, H. Waite, J. Bergervoet, K. Harish, Y. Zhang, C. Razzell, and R. Roovers, SiGe BiCMOS 1ns fast hopping frequency synthesizer for UWB radio, in *ISSCC Digest of Technical Papers*, San Francisco, Feb. 2005.
38. J. C. and M. Steyaert, Low-noise voltage controlled oscillators using enhanced LC-tanks, *IEEE Trans. Circuits Syst. II*, 42(12), 794–804, 1995.
39. B. Razavi, Analysis, modeling and simulation of phase noise in monolythic voltage controlled oscillators, in *Proceedings of the Custom Integrated Circuits Conference, CICC*, May 1995, pp. 323–326.
40. J. Craninckx and M. Steyaert, A 1.8GHz low phase noise voltage controlled oscillator with prescaler, *IEEE J. Solid-State Circuits*, 30(12), 1474–1482, 1995.
41. B. de Muer and M. Steyaert, *CMOS Fractional-N Synthesizers: Design for High Spectral Purity and Monolithic Integration*, Springer, Norwell, MA, 2003.
42. B. Gilbert, A precise four-quadrant multiplier with sub-nanosecond response, *IEEE J. Solid-State Circuits*, 3(4), 365–373, 1968.
43. A. Rofougaran, J. Y.-C. Chang, M. Rofougaran, S. Khorram, and A. A. Abidi, A 1 GHz CMOS RF front-end IC with wide dynamic range, in *Proceedings of the European Solid State Circuits Conference, ESSCIRC*, Sep. 1995, pp. 250–253.
44. M. Borremans and M. Steyaert, A 2V low distortion 1 GHz CMOS up-converter mixer, *IEEE J. Solid-State Circuits*, 33(3), 359–366, 1998.
45. M. Borremans, M. Steyaert, and T. Yoshitomi, A 1.5V wide band 3GHz CMOS quadrature direct up-converter for multi-mode wireless communications, in *Proceedings of the Custom Integrated Circuits Conference, CICC*, May 1998, pp. 79–82.
46. M. Pelgrom, A. Duinmaijer, and A. Welbers, Matching properties of MOS transistor, *IEEE J. Solid-State Circuits*, 24(5), 1433–1439, 1989.
47. D. Su and W. McFarland, A 2.5V, 1-W monolithic CMOS RF power amplifier, in *Proceedings of the Custom Integrated Circuits Conference, CICC*, May 1997, pp. 189–192.
48. K. C. Tsai and P. R. Gray, 1.9-GHz 1-W CMOS RF power amplifier for wireless communication, *IEEE J. Solid-State Circuits*, 34(7), 962–970, 1999.
49. C. Yoo and Q. Huang, A common-gate switched 0.9-W class-E power amplifier with 41% PAE in 0.25-μm CMOS, *IEEE J. Solid-State Circuits*, 36(5), 823–830, 2001.
50. T. Sowlati and D. Leenaerts, A 2.4GHz 0.18 μm CMOS self-biased cascode power amplifier with 23dBm output power, in *ISSCC Digest of Technical Papers*, San Francisco, Feb. 2002, pp. 294–295.

51. C. Fallesen and P. Asbeck, A 1W 0.35 μm CMOS power amplifier for GSM-1800 with 45% PAE, in *ISSCC Digest of Technical Papers*, San Francisco, Feb. 2001, pp. 158–159.
52. A. Shirvani, D. K. Su, and B. A. Wooley, A CMOS RF power amplifier with parallel amplification for efficient power control, in *ISSCC Digest of Technical Papers*, San Francisco, Feb. 2001.
53. V. R. Vathulya, T. Sowlati, and D. Leenaerts, Class 1 bluetooth power amplifier with 24 dBm output power and 48% PAE at 2.4GHz in 0.25 μm CMOS, in *Proceedings of the European Solid State Circuits Conference, ESSCIRC*, Villach, Austria, 2001.
54. I. Aoki, S. D. Kee, D. B. Rutledge, and A. Hajimiri, Fully integrated CMOS power amplifier design using the distributed active transformer architecture, *IEEE J. Solid-State Circuits*, 37(3), 1–13, 2002.
55. N. O. Sokal and A. D. Sokal, Class E—A new class of high-efficiency tuned single-ended switching power amplifiers, *IEEE Journal of Solid-State Circuits*, 10(3), June 1975, pp. 168–176.

5

PLL Circuits

5.1 Introduction **5**-1
What Is and Why Phase-Locked? • Basic Operation Concepts of PLLs • Classification of PLL Types
5.2 PLL Techniques **5**-2
Basic Topology • Loop Orders of the PLL • Tracking Process • Lock-In Process • Acquisition Process • Aided Acquisition • PLL Noise Performance
5.3 Building Blocks of PLL Circuit **5**-10
Voltage-Controlled Oscillators • Phase and Frequency Detectors • Loop Filters • Charge-Pump PLL • PLL Design Considerations
5.4 PLL Applications **5**-27
Clock and Data Recovery • Delay-Locked Loop • Frequency Synthesizer
Bibliography **5**-34

Muh-Tian Shiue
National Central University

Chorng-Kuang Wang
National Taiwan University

5.1 Introduction

5.1.1 What Is and Why Phase-Locked?

Phase-locked loop (PLL) is a circuit architecture that causes a particular system to track with another one. More precisely, PLL synchronizes a signal (usually a local oscillator output) with a reference or an input signal in frequency as well as in phase.

Phase locking is a useful technique that can provide effective synchronization solutions in many data transmission systems such as optical communications, telecommunications, disk drive systems, and local networks, in which data are transmitted in baseband or passband. In general, only data signals are transmitted in most of these applications, namely, clock signals are not transmitted in order to save hardware cost. Therefore, the receiver should have some mechanisms to extract the clock information from the received data stream in order to recover the transmitted data. The scheme is called a timing recovery or clock recovery.

The cost of electronic interfaces in communication systems raises as the data rate gets higher. Hence, high-speed circuits are the critical issue of the high data rate systems implementation, and the advanced very large scale integration (VLSI) technology plays an important role in cost reduction for the high-speed communication systems.

5.1.2 Basic Operation Concepts of PLLs

Typically, as shown in Figure 5.1, a PLL consists of three basic functional blocks: a phase detector (PD), a loop filter (LF), and a voltage-controlled oscillator (VCO). PD detects the phase difference between the VCO output and the input signal, and generates a signal proportional to the phase error. The PD output

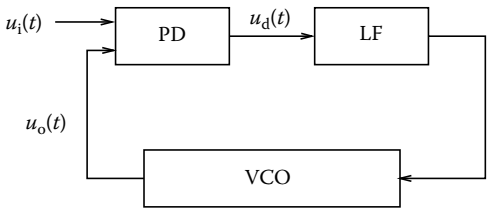


FIGURE 5.1 Basic block diagram of the PLL.

contains a direct current (DC) component and an alternative current (AC) component, the former is accumulated and the latter is filtered out by the LF. The LF output that is near a DC signal is applied to the VCO. This almost DC control voltage changes the VCO frequency toward a direction to reduce the phase error between the input signal and the VCO. Depending on the type of LF used, the steady-state phase error will be reduced to zero or to a finite value.

PLL has an important feature, which is the ability to suppress both the noises superimposed on the input signal and generated by the VCO. In general, the more narrow bandwidth the PLL has, the more effectively the filtering of the superimposed noises can be achieved. Although a narrow bandwidth is better for rejecting large amounts of the input noise, it also prolongs the settling time in the acquisition process. Then, the error of the VCO frequency cannot be reduced rapidly. So there is a trade-off between jitter filtering and fast acquisition.

5.1.3 Classification of PLL Types

Different PLL types have been built from different classes of building blocks. The first PLL integrated circuit (IC) appeared around 1965 and consisted of purely analog devices. In the so-called linear PLLs (LPLLs), an analog multiplier (four-quadrant) is used as the PD, the LF is built of a passive or an active RC filter, and the VCO is used to generate the output signal of the PLL. In most cases, the input signal to this LPLL is a sine wave, whereas the VCO output signal is a symmetrical square wave.

The classical digital PLL (DPLL) uses a digital PD such as an exclusive OR (XOR) gate, a JK-flipflop, or a phase-frequency detector (PFD). The remaining blocks are still the same as LPLL. In many aspects, the DPLL performance is similar to the LPLL.

The function blocks of the all digital PLL (ADPLL) are implemented by purely digital circuits, and the signals within the loop are digital, too. Digital versions of the PD are the same as DPLL. The digital LF is built of an ordinary up/down counter, N-before-M counter or K-counter [1]. The digital counterpart of the VCO is the digital-controlled oscillator [2,3].

In analogy to filter designs, PLLs can be implemented by software such as a microcontroller, microcomputer, or digital signal processing (DSP); this type of PLL is called software PLL (SPLL).

5.2 PLL Techniques

5.2.1 Basic Topology

A PLL is a feedback system that operates and minimizes the phase difference between two signals. Referring to the basic function block diagram of a PLL as shown in Figure 5.1, it typically consists of a PD, an LF, and a VCO. The PD works as a phase error detector and an amplifier. It compares the phase of the VCO output signal $u_o(t)$ with the phase of the reference signal $u_i(t)$ and develops an output signal $u_d(t)$ that is proportional to the phase error θ_e. Within a limited range, the output signal can be expressed as

$$u_d(t) = k_d \theta_e \tag{5.1}$$

where k_d with the unit of V/rad represents the gain of the PD.

The output signal $u_d(t)$ of the PD consists of a DC component and a superimposed AC component. The latter is undesired and removed by the LF. In general, the LF is a low-pass filter (LPF) to generate an almost DC control voltage for the VCO to oscillate at the frequency equal to the input frequency.

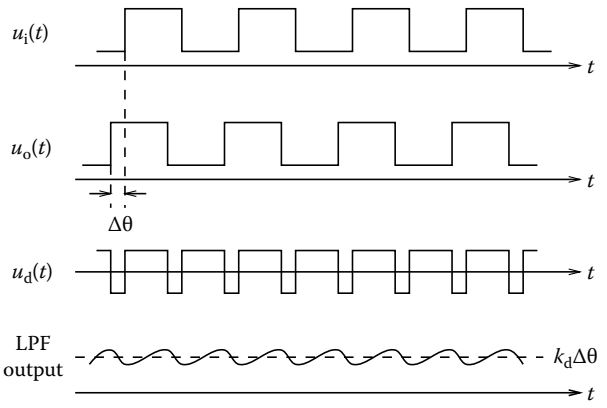


FIGURE 5.2 Waveforms in a PLL.

How the building blocks of a basic PLL work together will be explained below. At first, assume both the waveforms of input signal and VCO output are rectangular. Furthermore, it is assumed that the angular frequency ω_i of the input signal $u_i(t)$ is equivalent to the central frequency ω_o of the VCO signal $u_o(t)$. Now a small positive frequency step is applied to $u_i(t)$ at $t = t_0$ (shown in Figure 5.2). $u_i(t)$ accumulates the phase increments faster than $u_o(t)$ of VCO does. If the PD can response wider pulses increasingly, a higher DC voltage is accordingly generated at the LF output to increase the VCO frequency. Depending on the type of the LF that will be discussed later, the final phase error will be reduced to zero or a finite value.

It is important to note from the descriptions above that the loop locks only after the two conditions are satisfied: (1) ω_i and ω_o are equal and (2) the phase difference between the input $u_i(t)$ and the VCO output $u_o(t)$ settles to a steady-state value. If the phase error varies with time so fast that the loop is unlocked, the loop must keep on the transient process, which involves both "frequency acquisition" and "phase acquisition."

To design a practical PLL system, it is required to know the status of the responses of the loop if (1) the input frequency is varied slowly (tracking process), (2) the input frequency is varied abruptly (lock-in process), and (3) the input and the output frequencies are not equal initially (acquisition process). Using LPLL as an example, these responses will be shown in Sections 5.2.3 through 5.2.5.

5.2.2 Loop Orders of the PLL

Figure 5.3 shows the linear model of a PLL. According to the control theory, the closed-loop transfer function of PLL can be derived as

$$H(s) \triangleq \frac{\theta_o(s)}{\theta_i(s)} = \frac{k_d k_o F(s)}{s + k_d k_o F(s)} \tag{5.2}$$

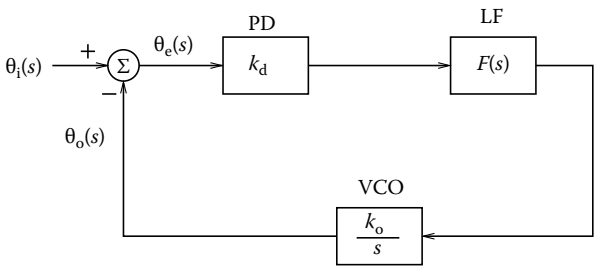


FIGURE 5.3 Linear model of PLL.

where

k_d with units V/rad is called the PD gain

k_o is the VCO gain factor and has units rad/(s-V)

In addition to the phase transfer function, a phase-error transfer function $H_e(s)$ is derived as follows:

$$H_e(s) \triangleq \frac{\theta_e(s)}{\theta_i(s)} = \frac{s}{s + k_d k_o F(s)} \tag{5.3}$$

The loop order of the PLL depends on the characteristics of the LF. Therefore the LF is a key component that affects the PLL dynamic behavior. A PLL with an LF consisted of simple amplifier or attenuator is called a first-order PLL. As shown in Figure 5.3, set $F(s) = 1$ and the closed-loop transfer function can be derived as

$$H(s) = \frac{k}{s + k} \tag{5.4}$$

where the DC loop gain $k = k_d k_o$. If fast tracking is required, a high DC loop gain k is necessary for the bandwidth of the PLL being wide enough because the DC loop gain k is the only parameter available. Such a design is not suitable for noise suppression. Therefore, fast tracking and narrow bandwidth are incompatible in a first-order loop.

A commonly used LF is the passive lag filter. The transfer function is

$$F(s) = \frac{1}{1 + s\tau} \tag{5.5}$$

The closed-loop transfer function can be derived as

$$H(s) = \frac{k_d k_o/\tau}{s^2 + (1/\tau)s + k_d k_o/\tau} = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \tag{5.6}$$

where

$\omega_n = \sqrt{\frac{k_d k_o}{\tau}}$ is the "natural frequency"

$\zeta = \frac{1}{2}\sqrt{\frac{1}{\tau k_d k_o}}\,\omega_n$ is the "damping factor"

These two parameters are important to characterize a PLL. Now, a second-order PLL is obtained and there are two parameters (τ and $k = k_o k_d$) available to achieve fast tracking as well as the noise suppression. Then three loop parameters (ω_n, ζ, k) must be determined. In addition, the phase-error transfer function $H_e(s)$ can be further derived as follows:

$$H_e(s) = \frac{s(s + 1/\tau)}{s^2 + (1/\tau)s + k_d k_o/\tau} \tag{5.7}$$

5.2.3 Tracking Process

The linear model of a PLL shown in Figure 5.3 is suitable for analyzing the tracking performance of a PLL that is almost in lock, that is, only with a small phase error. If the phase error changes too abruptly, the PLL fails to lock, and a large phase error is induced even though the change happens only momentarily. The unlock condition is a nonlinear process that cannot be analyzed via the linear model. The acquisition process will be described in Section 5.2.5.

At first, consider that a step phase error expressed as $\theta_i(t) = \Delta\theta u(t)$ is applied to the input. The Laplace transform of the input is $\theta_i(s) = \frac{\Delta\theta}{s}$ that is substituted into Equation 5.7 to get

$$\theta_e(s) = \frac{\Delta\theta}{s} \frac{s(s + 1/\tau)}{s^2 + (1/\tau)s + k_d k_o/\tau} \tag{5.8}$$

According to the final value theorem of the Laplace transform,

$$\lim_{t\to\infty} \theta_e(t) = \lim_{s\to 0} s\theta_e(s) = \frac{\Delta\theta}{k_d k_o}$$

In another word, the loop will eventually track on the step phase change with a steady-state phase error proportional to the DC loop gain. If it is necessary to have a high DC loop gain in order to reduce the steady-state phase error and a very narrow bandwidth for improving the noise suppression, the loop will be severely underdamped and the transient response will be poor.

When a step change of frequency $\Delta\omega$ is applied to the input, the input phase change is a ramp, that is, $\theta_i(t) = \Delta\omega t$, therefore $\theta_i(s) = \frac{\Delta\omega}{s^2}$. Substituting $\theta_i(s)$ in Equation 5.7 and applying the final value theorem, then

$$\begin{aligned} \theta_v &= \lim_{t\to\infty} \theta_e(t) = \lim_{s\to 0} s\theta_e(s) \\ &= \lim_{s\to 0} \frac{\Delta\omega}{s^2} \frac{s(s + 1/\tau)}{s^2 + (1/\tau)s + (k_d k_o/\tau)} \end{aligned} \tag{5.9}$$

where θ_v is called the "velocity error" or "static phase error" [4]. In practice, the input frequency almost never agrees exactly with the VCO free-running frequency, that is, usually there is a frequency difference $\Delta\omega$ between the two. From Equation 5.9, the velocity error will be infinite while there is a frequency difference $\Delta\omega$. Another commonly used LF is the active lead-lag LF with the transfer function $F(s)$ described as follows:

$$F(s) = k_a \frac{1 + s\tau_2}{1 + s\tau_1} \tag{5.10}$$

where τ_1, τ_2, and k_a are the two time constants and DC gain of an active lead-lag LF, respectively. Substituting $\theta_i(s)$ in Equation 5.3 and applying the final value theorem, then

$$\begin{aligned} \theta_v &= \lim_{s\to 0} s\theta_e(s) = \lim_{s\to 0} \frac{\Delta\omega}{s + k_d k_o F(s)} \\ &= \frac{\Delta\omega}{k_d k_o F(0)} = \frac{\Delta\omega}{k_v} \end{aligned} \tag{5.11}$$

From Equation 5.11, if the PLL has a high DC loop gain, that is, $k_d k_o F(0) \gg \Delta\omega$, the steady-state phase error corresponding to a step frequency error input approaches to zero. This is the reason that a high gain loop has a good tracking performance. Now the advantage of a second-order loop using an active LF with high DC gain is evident. The active lead-lag LF with a high DC gain will make the steady-state phase error approach to zero and the noise bandwidth be narrow simultaneously, which is impossible in a first-order loop.

If the input frequency is changed linearly with time at a rate of $\Delta\omega$, that is $\theta_i(t) = \frac{1}{2}\Delta\omega\ t^2$, $\theta_i(s) = \frac{\Delta\omega}{s^3}$. Using an active LF with high DC gain and applying the final value theorem of Laplace transform, it is derived that

$$\theta_a = \lim_{t\to\infty} \theta_e(t) = \lim_{s\to 0} s\theta_e(s) = \frac{\Delta\omega}{\omega_n^2} \tag{5.12}$$

where θ_a is called an "acceleration error" (sometimes calls "dynamic tracking error" or "dynamic lag") [4].

In some applications, PLL needs to track an accelerating phase error without static tracking error. When frequency ramp is applied, the static phase error will be

$$\theta_e(s) = \lim_{s \to 0} \frac{\Delta\omega}{s(s + k_d k_o F(s))} \tag{5.13}$$

In order to have θ_e zero, it is necessary to make $F(s)$ be a form of $\frac{G(s)}{s^2}$, where $G(0) \neq 0$. $\frac{G(s)}{s^2}$ implies that the LF has two cascade integrators. This results in a third-order loop. In order to eliminate the static acceleration error, a third-order loop is very useful for some special applications such as satellite and missile systems.

Based on Equation 5.12, a large natural frequency ω_n is used to reduce the static tracking phase error in a second-order loop; however, a wide natural frequency has an undesired noise filtering performance. In the contrast, the zero tracking phase error for a frequency ramp error is concordant with a small loop bandwidth in a third-order loop. In practice, there are three basic types of LF: passive lead-lag filter, active lead-lag filter, and active proportional and integral (PI) filter. The characteristics of the three types of LF and their effects on the PLL will be described in Section 5.3.3. Besides, a high-order filter is used for critical applications because it provides better noise filtering, initial acquisition, and fast tracking. However it is difficult to design a high-order loop due to some problems such as loop stability.

All the preceding analysis on the tracking process is under the assumption that the phase error is relatively small and the loop is linear. If the phase error is large enough to make the loop drop out of lock, the linear assumption is invalid. For a sinusoidal-characteristic PD, the exact phase expression of Equation 5.11 should be

$$\sin\theta_v = \frac{\Delta\omega}{k_v} \tag{5.14}$$

The sine function has solutions only when $\Delta\omega \leq k_v$. However, there is no solution if $\Delta\omega > k_v$. This is the case the loop loses lock and the output of the PD will be beat notes signal rather than a DC control voltage. Therefore, k_v can be used to define the ''hold range'' of the PLL, that is

$$\Delta\omega_H = \pm k_v = k_o k_d F(0) \tag{5.15}$$

The hold range is the frequency range in which a PLL is able to maintain lock ''statically.'' Namely, if input frequency offset exceeds the hold range statically, the steady-state phase error would drop out of the linear range of the PD and the loop loses lock. k_v is the function of k_o, k_d, and $F(0)$. The DC gain $F(0)$ of the LF depends on the filter type. Therefore, it is important to make an LF have a high DC gain for extending the hold range. Referring to the characteristics of the three basic types of LF described in Section 5.3.3, the hold range $\Delta\omega_H$ can be $k_o k_d$, $k_o k_d k_a$, and ∞ for passive lead-lag filter, active lead-lag filter, and active PI filter, respectively. The hold range expressed in Equation 5.15 is not correct when some other components in PLL are saturated earlier than the PD. When the PI filter is used, the real hold range is actually determined by the control range of the VCO.

Considering the dynamic phase error θ_a in a second-order loop, the exact expression for a sinusoidal characteristic PD is

$$\sin\theta_a = \frac{\Delta\omega}{\omega_n^2} \tag{5.16}$$

which implies that the maximum change rate of the input frequency is ω_n^2. If the rate exceeds ω_n^2, the loop will fall out of lock.

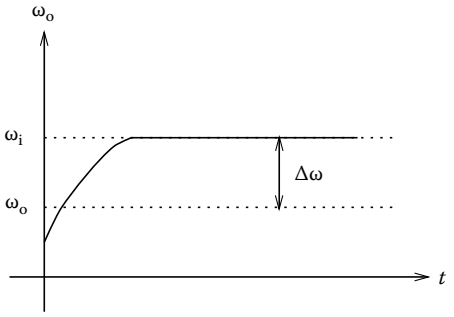


FIGURE 5.4 Lock-in process of the PLL.

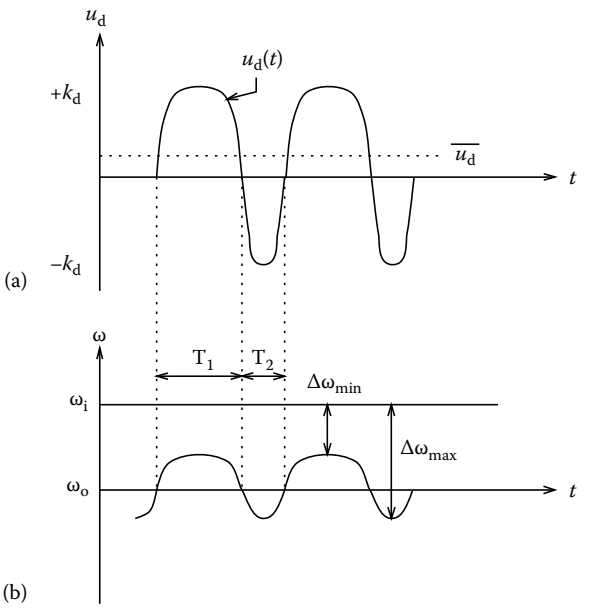


FIGURE 5.5 Pull-in process of the PLL.

5.2.4 Lock-In Process

The "lock-in" process is defined as PLL locks within one single beat note between the input and the output (VCO output) frequency. The maximum frequency difference between the input and the output that PLL can lock within one single beat note is called the "lock-in range" of the PLL.

Figure 5.4 shows a case of PLL lock-in process that a frequency offset $\Delta\omega$ is less than the lock-in range, and the lock-in process happens. Then PLL will lock within one single beat note between ω_i and ω_o. In Figure 5.5b, the frequency offset $\Delta\omega$ between input (ω_i) and output (ω_o) is larger than the lock-in range, hence the lock-in process will not take place, at least not instantaneously.

Suppose the PLL is unlocked initially. The input frequency ω_i is $\omega_o + \Delta\omega$. If the input signal $v_i(t)$ is a sine wave and given by

$$v_i(t) = A_i \sin(\omega_o t + \Delta\omega t) \tag{5.17}$$

And the VCO output signal $v_o(t)$ is usually a square wave written as a Walsh function [5]

$$v_o(t) = A_o W(\omega_o t) \tag{5.18}$$

$v_o(t)$ can be replaced by the Fourier series,

$$v_o(t) = A_o \left[\frac{4}{\pi}\cos(\omega_o t) + \frac{4}{3\pi}\cos(3\omega_o t) + \cdots\right] \tag{5.19}$$

So the PD output v_d is

$$\begin{aligned} v_d(t) = v_i(t)v_o(t) &= A_i A_o \left[\frac{2}{\pi}\sin(\Delta\omega t) + \cdots\right] \\ &= k_d \sin(\Delta\omega t) + \text{high} - \text{frequency terms} \end{aligned} \tag{5.20}$$

The high frequency components can be filtered out by the LF. The output of the LF is given by

$$v_f(t) \approx k_d |F(\Delta\omega)| \sin(\Delta\omega t) \tag{5.21}$$

The peak frequency deviation based on Equation 5.21 is equal to $k_d k_o|F(\Delta\omega)|$. If the peak deviation is larger than the frequency error between ω_i and ω_o, the lock-in process will take place. Hence the lock-in range is consequently given by

$$\Delta\omega_L = k_d k_o |F(\Delta\omega_L)| \tag{5.22}$$

The lock-in range is always larger than the corner frequency $\frac{1}{\tau_1}$ and $\frac{1}{\tau_2}$ of the LF in practical cases. An approximation of the LF gain $F(\Delta\omega_L)$ is shown as follows:

For the passive lead-lag filter

$$F(\Delta\omega_L) \approx \frac{\tau_2}{\tau_1 + \tau_2}$$

For the active lead-lag filter

$$F(\Delta\omega_L) \approx k_a \frac{\tau_2}{\tau_1}$$

For the active PI filter

$$F(\Delta\omega_L) \approx \frac{\tau_2}{\tau_1}$$

τ_2 is usually much smaller than τ_1, the $F(\Delta\omega_L)$ can be further approximated as follows:

For the passive lead-lag filter

$$F(\Delta\omega_L) \approx \frac{\tau_2}{\tau_1 + \tau_2}$$

For the active lead-lag filter

$$F(\Delta\omega_L) \approx k_a \frac{\tau_2}{\tau_1}$$

For the active PI filter

$$F(\Delta\omega_L) \approx \frac{\tau_2}{\tau_1}$$

Substituting above equations in Equation 5.22 and assuming a high gain loop,

$$\Delta\omega_L \approx 2\zeta\omega_n \tag{5.23}$$

can be gotten for all three types of LF shown in Figure 5.12.

5.2.5 Acquisition Process

Suppose that the PLL does not lock initially, the input frequency is $\omega_i = \omega_o + \Delta\omega$, where ω_o is the initial frequency of VCO. If the frequency error $\Delta\omega$ is larger than the lock-in range, the lock-in process will not happen. Consequently the output signal $u_d(t)$ of the PD shown in Figure 5.5a is a sine wave that has the frequency of $\Delta\omega$. The AC PD output signal $u_d(t)$ passes through the LF. Then the output $u_f(t)$ of the LF modulates the VCO frequency. As shown in Figure 5.5b, when ω_o increases, the frequency difference between ω_i and ω_o becomes smaller and vice versa. Therefore, the PD output $u_d(t)$ becomes asymmetric when the duration of positive half-periods of the PD output is larger than the negative ones. The average value $\overline{u_d(t)}$ of the PD output therefore goes positive slightly. Then the frequency of VCO will be pulled up until it reaches the input frequency. This phenomenon is called a ''pull-in process.''

Because the pull-in process is a nonlinear behavior, the mathematical analysis is quite complicated. According to the results of [1], the pull-in range and the pull-in time depend on the type of LF. For an active lead-lag filter with a high gain loop, the pull-in range is

$$\Delta\omega_P \approx \frac{4\sqrt{2}}{\pi}\sqrt{\zeta\omega_n k_o k_d} \tag{5.24}$$

and the pull-in time is

$$T_P \approx \frac{\pi^2}{16}\frac{\Delta\omega_0^2 k_a}{\zeta\omega_n^3} \tag{5.25}$$

where $\Delta\omega_0$ is the initial frequency error. Equations 5.24 and 5.25 should be modified for different types of PDs [1].

5.2.6 Aided Acquisition

The PLL bandwidth is always too narrow to lock a signal with large frequency error. Furthermore, the frequency acquisition is slow and impractical. Therefore, there are aided frequency-acquisition techniques to solve this problem such as the frequency locked-loop (FLL) and the bandwidth-widening methods.

The FLL, which is very much similar to a PLL, is composed of a frequency discriminator, an LF, and a VCO. PLL is a coherent mechanism to recover a signal buried in noise. An FLL, in contrast, is a noncoherent scheme that cannot distinguish the phase error between input signal and VCO signal. Therefore an FLL can only be useful to provide the signal frequency which exactly synchronizes with the reference frequency source.

The major difference between PLL and FLL is the PD and the frequency discriminator. The frequency discriminator is the frequency detector in the FLL. It generates a voltage proportional to the frequency difference between the input and the VCO. The frequency difference will be driven to zero in a negative feedback fashion. If a linear frequency detector is employed, it can be shown that the frequency-acquisition time is proportional to the logarithm of the frequency error [6]. In the literature, some frequency detectors-like quadricorrelator [7], balance quadricorrelator [8], rotational frequency detector [9], and frequency delimiter [10] are disclosed.

5.2.7 PLL Noise Performance

In high-speed data recovery applications, a better performance of the VCO and the overall PLL itself is desired. In a consequence, the random variations of the sampling clock, so-called jitter, is the critical performance parameter.

Jitter sources of PLL in the case of using a ring VCO mainly come from the input and the VCO itself. The ring oscillator jitter is associated with the power supply noise, the substrate noise, $1/f$ noise, and the thermal noise. The former two noise sources can be reduced by fully differential circuit structure. $1/f$ noise, on the other hand, can be rejected by the tracking capability of the PLL. Therefore, the thermal noise is the worst noise source. From the analysis of [18], the one stage RMS timing jitter error of the ring oscillator normalized to the time delay per stage can be shown as

$$\frac{\Delta\tau_{rms}}{t_d} \approx \sqrt{\frac{2KT}{C_L}}\left(\sqrt{1+\frac{2}{3}a_v}\right)\frac{1}{V_{pp}} \tag{5.26}$$

where

C_L is the load capacitance

$\sqrt{1+\frac{2}{3}a_v}$ is called the noise contribution factor ζ

a_v is the small-signal gain of the delay cell

V_{pp} is the VCO output swing

From Equation 5.26, for a fixed output bandwidth, higher gain contributes larger noise.

Because the ring oscillator is a feedback architecture, the noise contribution of a single delay cell may be amplified and filtered by the following stage. To consider two successive stages, Equation 5.26 can be rearranged as [18]

$$\frac{\Delta\tau_{rms}}{t_d} \approx \sqrt{\frac{2KT}{C_L}}\frac{1}{(V_{gs}-V_t)}\zeta \tag{5.27}$$

Therefore, the cycle-to-cycle jitter of the ring oscillator in a PLL can be predicted by [18]

$$\overline{(\Delta\tau_N)^2} = \frac{KT}{I_{ss}}\frac{a_v\zeta^2}{(V_{gs}-V_t)}T_o \tag{5.28}$$

where

I_{ss} is the rail current of the delay cell

T_o is the output period of the VCO

Based on Equation 5.28, designing a low jitter VCO ($V_{gs}-V_t$) should be as large as possible. For fixed delay and fixed current, a lower gain of each stage is better for jitter performance, but the loop gain must satisfy the Barkhausen criterion. From the viewpoint of VCO jitter, a wide bandwidth of PLL can correct the timing error of the VCO rapidly [14]. If the bandwidth is too wide, the input noise jitter may be so large that dominates the jitter performance of the PLL. Actually this is a trade-off.

For a PLL design, the natural frequency and the damping factor are the key parameters to be determined by designers. If the input signal-to-noise ratio $(SNR)_i$ is defined, then the output signal-to-noise ratio $(SNR)_o$ can be obtained [4]

$$(SNR)_o = (SNR)_i\frac{B_i}{2B_L} \tag{5.29}$$

where

B_i is the bandwidth of the prefilter

B_L is the noise bandwidth

Hence the B_L can be derived using Equation 5.29. And the relationship of B_L with ω_n and ζ is

$$B_L = \frac{\omega_n}{2}\left(\zeta+\frac{1}{4\zeta}\right) \tag{5.30}$$

Therefore the ω_n and ζ can be designed to satisfy the $(SNR)_o$ requirement.

Besides the system and the circuit designs, jitter can be reduced in the board level design. Board jitter can be alleviated by better layout and noise decoupling schemes such as appending proper decouple and bypass capacitances.

5.3 Building Blocks of PLL Circuit

5.3.1 Voltage-Controlled Oscillators

The function of a VCO is to generate a stable and periodic waveform whose frequency can be varied by an applied control voltage. The relationship between the control voltage and the oscillation frequency

depends upon the circuit architecture. A linear characteristic is generally preferred because of its wider applications. As a general classification, VCO can be categorized roughly into two types by the output waveforms: (1) harmonic oscillators that generate nearly sinusoidal outputs and (2) relaxation oscillators that provide square or triangle outputs.

In general, a harmonic oscillator is composed of an amplifier that provides an adequate gain and a frequency-selective network that feeds a certain output frequency range back to the input. LC tank oscillators and crystal oscillators belong to this type. Generally, the harmonic oscillators have the following advantages: (1) superior frequency stability while the conditions of temperature, power supply, and noise are included; and (2) good frequency accuracy control due to that the oscillation frequency is determined by a tank circuit or a crystal.

Essentially, harmonic oscillators are not compatible with monolithic IC technology and their frequency tuning range is limited. On the contrary, relaxation oscillators are easy to be implemented in monolithic ICs. Since frequency is normally proportional to a controlled-current or -voltage and inversely proportional to timing capacitors, the frequency of oscillation can be varied linearly over a very wide range. On the other hand, the ease of frequency tuning brings in drawbacks, such as poor frequency stability and frequency inaccuracy.

Relaxation oscillators are the most commonly used oscillator configuration in monolithic IC design because they can operate in a wide frequency range with a minimum number of external components. According to the mechanism of the oscillator topology employed, relaxation oscillators can be further categorized into three types: (1) grounded capacitor VCO [20], (2) emitter-coupled VCO, and (3) delay-based ring VCO [21]. The operation of the first two oscillators are similar in the sense that the time duration spent in each state is determined by the timing components and the charge/discharge currents. The ring oscillator is one of the relaxation oscillators and has received considerable attentions recently in high frequency PLL applications for clock synchronization and timing recovery. Because the ring oscillator can provide high frequency oscillation with simple digital-like circuits that are compatible with digital technology, it is suitable for VLSI implementations.

In order to achieve high rejection of power supply and substrate noises, both the signal path and the control path of a VCO must be fully differential. A common ring oscillator topology in monolithic PLLs is shown in Figure 5.6. The delay-based ring VCO operates quite differently since the timing relies on the delay in each gain stages that are connected in a ring configuration. The loop oscillates with a period equal to $2NT_d$, where T_d is the delay of each stage. The oscillation can be obtained when the total phase shift is zero and the loop gain is greater or equal to unity at a certain frequency. To vary the frequency of oscillation, the effective number of stages or the delay of each stage must be changed. The first approach is called ''delay interpolating'' VCO [21], where a shorter delay path and a longer delay path are used in parallel. The total delay is tuned by increasing the gain of one path and decreasing the other, and the total delay is a weighted sum of the two delay paths. The second approach is to vary the delay time of each stage to adjust the oscillation frequency. The delay of each stage is tuned by varying the capacitance or the resistance seen at the output node of each stage.

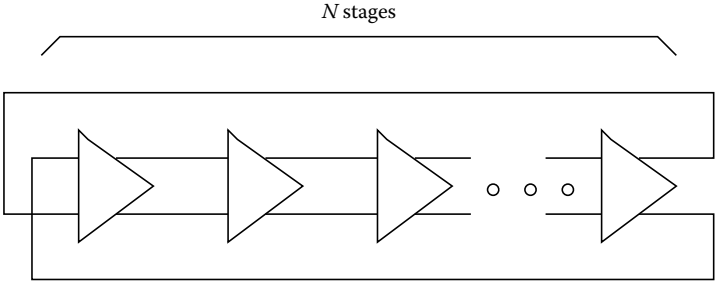


FIGURE 5.6 Ring oscillator.

Because the tuning range of the capacitor is small and the maximum oscillation frequency is limited by the minimum value of the load capacitor, the ''resistive tuning'' is a better alternative technique. Resistive tuning method provides a large, uniform frequency tuning range and leads itself easily to a differential control. In Figure 5.7a, the on-resistance of the triode PMOS loads is adjusted by V_{cont}. The more V_{cont} decreases, the more the delay of the stage drops because the time constant at the output node is decreased. However, the small-signal gain decreases as well when V_{cont} decreases. The circuit eventually fails to oscillate when the loop gain at the oscillation frequency is less than unity. In Figure 5.7b, the delay of gain stage is tuned by adjusting the tail current, but the small-signal gain remains constant. So the circuit is better than Figure 5.7a. As shown in Figure 5.7c [22], the PMOS current source with a pair of cross-coupled diode loads provides a differential load impedance that is independent of common-mode voltage. This makes the cell delay insensitive to common-mode noise. Figure 5.7d is a poor delay cell for a ring oscillator because the tuning range is very small.

The minimum number of stages that can be used while maintaining a reliable operation is an important issue in a ring oscillator design. When the number of stages decreases, the required phase

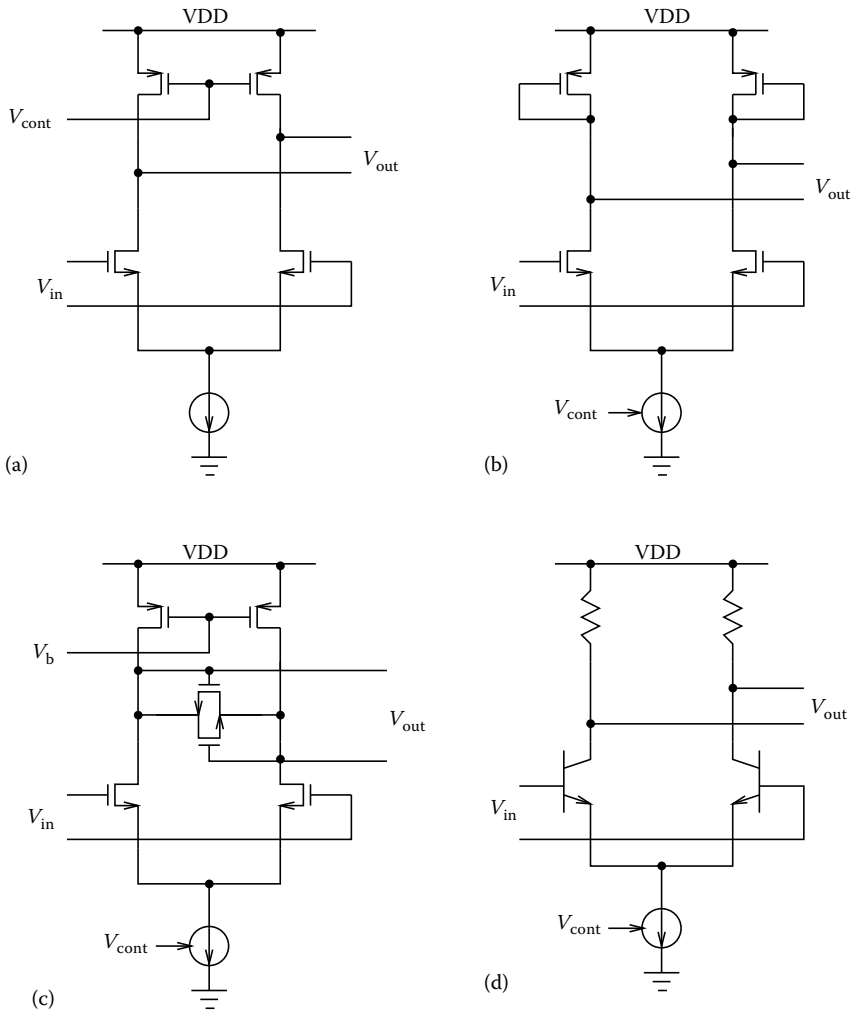


FIGURE 5.7 The gain stages using of resistive tuning.

shift and DC gain per stage increases. Two-stage bipolar ring oscillator can be designed reliably [23], but CMOS implementations are not. Thus, CMOS ring oscillators utilize three or more stages typically.

5.3.2 Phase and Frequency Detectors

The PD type has the influence on the dynamic range of PLLs. Hold range, lock-in range, and pull-in range are analyzed in Section 5.2 based on the multiplier PD. Most of the other types of PD have a greater linear output span and a larger maximum output swing than a sinusoidal characteristic PD. A larger tracking range and a larger lock limit are available if the linear output range of PD increases. The three widely used PDs are XOR PD, edge-triggered JK-flipflop, and PFD. The characteristics of these PDs are plotted in Figure 5.8.

The XOR PD can maintain phase tracking when the phase error θ_e is confined in the range of

$$\frac{-\pi}{2} < \theta_e < \frac{\pi}{2}$$

as shown in Figure 5.8a. The zero phase error takes place when the input signal and the VCO output are quadrature in phase as shown in Figure 5.9a. As the phase difference deviates from $\frac{\pi}{2}$, the output duty cycle is no longer 50%, which provides a DC value proportional to the phase difference as shown in Figure 5.9b. But the XOR PD has a steady-state phase error if the input signal or the VCO output are asymmetric.

The JK-flipflop PD shown in Figure 5.10, also called a two-state PD, is barely influenced by the asymmetric waveform because it is edge-triggered. The zero phase error happens when the input signal and the VCO output are out-off phase as illustrated in Figure 5.10a. As shown in Figure 5.8b, the JK-flipflop PD can maintain phase tracking when the phase error is within the range of

$$-\pi < \theta_e < \pi$$

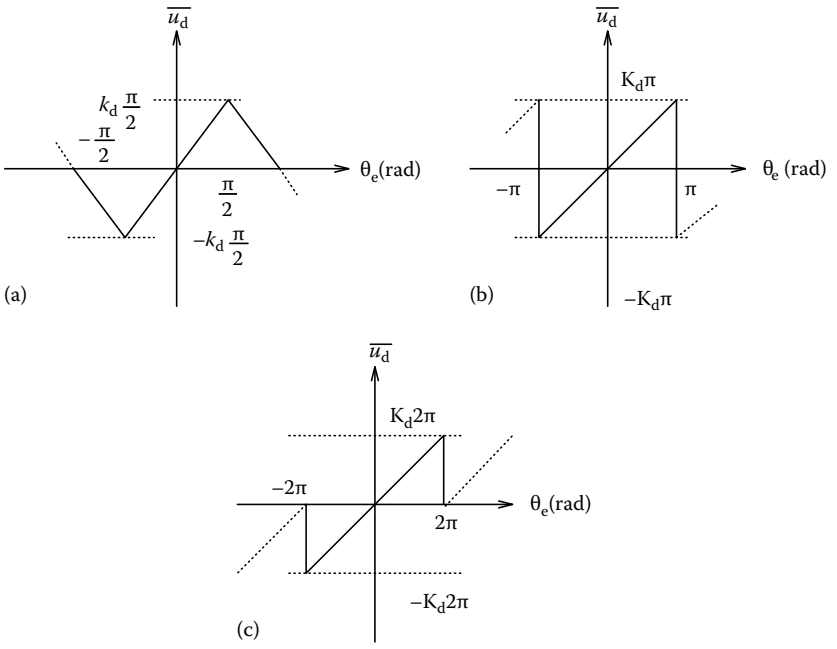


FIGURE 5.8 PD characteristics of (a) XOR, (b) JK-flipflop, and (c) PFD.

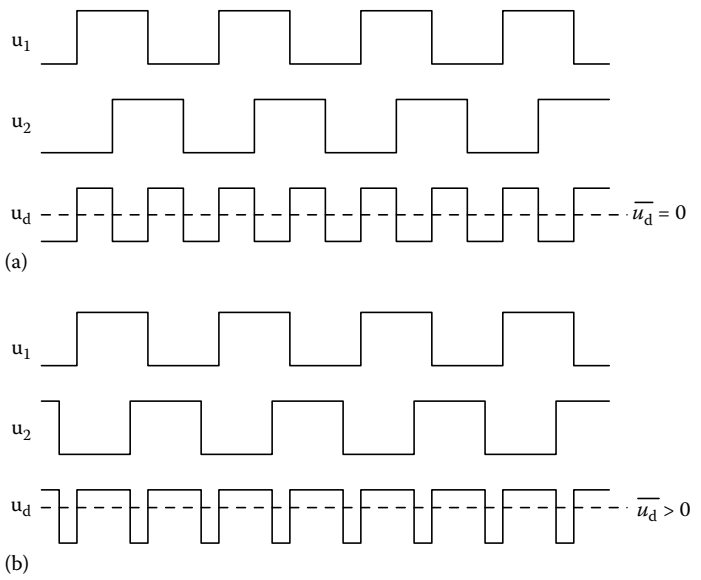


FIGURE 5.9 Waveforms of the signals for the XOR PD: (a) waveforms at zero phase error and (b) waveforms at positive phase error.

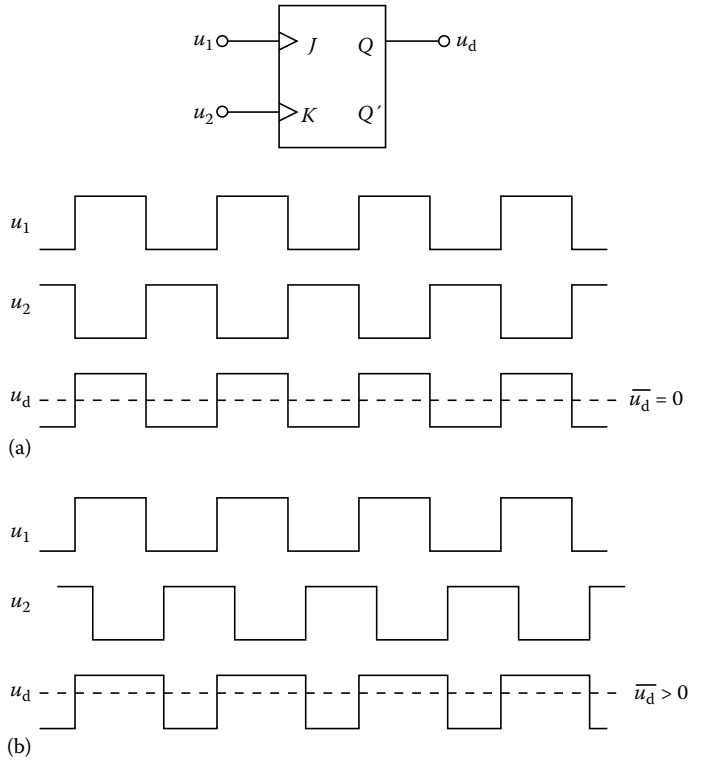


FIGURE 5.10 Waveforms of the signals for the JK-flipflop PD: (a) waveforms at zero phase error and (b) waveforms at positive phase error.

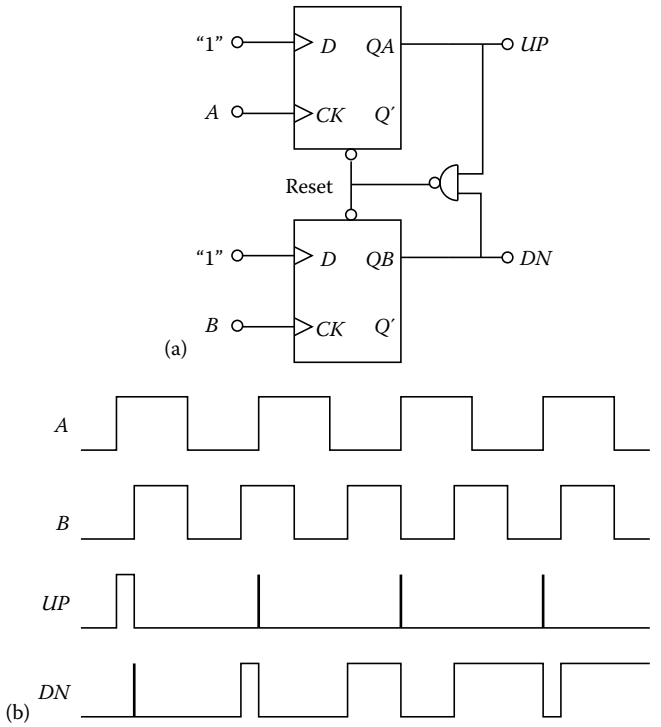


FIGURE 5.11 (a) PFD diagram and (b) inputs and outputs waveforms of PFD.

Here, a positive edge appearing at the J input triggers the flipflop into "high" state ($Q = 1$), and the rising edge of u_2 drives Q to zero. Figure 5.10b shows the output waveforms of the JK-flipflop PD for $\theta_e > 0$.

The PFD output depends not only on the phase error, but also on the frequency error. The characteristic is shown in Figure 5.8c. When the phase error is greater than 2π, the PFD works as a frequency detector. The operation of a typical PFD is as follows and the waveforms is shown in Figure 5.11. If the frequency of input A, ω_A, is less than the frequency of input B, ω_B, then the total width of positive pulses appearing at Q_A is shorter than at Q_B. Conversely, if $\omega_A > \omega_B$, the total width of positive pulses appears at Q_A is longer than at Q_B. If $\omega_A = \omega_B$, then the PFD generates pulses at either Q_A or Q_B with a width equal to the phase difference between the two inputs. The outputs Q_A and Q_B are usually called the "up" and "down" signals, respectively. If the input signal fails, which usually happens at the non-return-to-zero (NRZ) data recovery applications during missing or extra transmissions, the output of the PFD would stick on the high state (or low state). This condition may cause VCO to oscillate fast or slow abruptly, which results in noise jitter or even losing lock. This problem can be remedied by additional control logic circuits to make the PFD output to toggle back and forth between the two logic level with 50% duty cycle [19], the loop is interpreted as zero phase error. The "rotational FD" described by Messerschmitt can also solve this issue [9]. The output of a PFD can be converted to a DC control voltage by driving a three-state charge-pump which will be described in Section 5.3.4.

5.3.3 Loop Filters

For a PLL with given VCO and PD, LF is the critical component to determine the PLL characteristics, such as the damping factor that determines the relative stability, open-loop/closed-loop bandwidths that relate to the convergence speed in the initial state and the tracking capability in the steady state, and so on. Various types of LF will be introduced in this section.

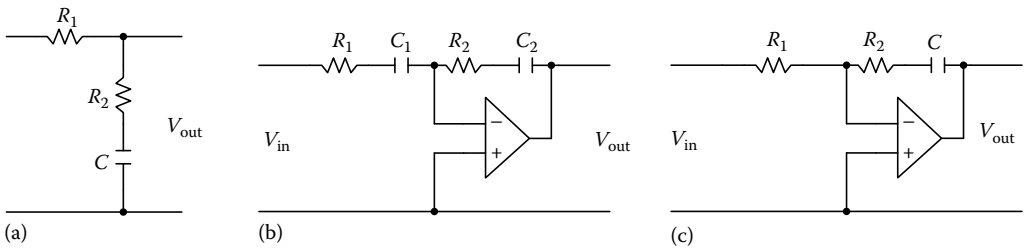


FIGURE 5.12 (a) Passive lead-lag filter, (b) active lead-lag filter, and (c) active PI filter.

5.3.3.1 Continuous-Time LFs

Figure 5.12 shows three types of LF that are widely used. Figure 5.12a is a passive lead-lag filter with transfer function $F(s)$ given by

$$F(s) = \frac{1 + s\tau_2}{1 + s(\tau_1 + \tau_2)} \tag{5.31}$$

where

$\tau_1 = R_1 C$
$\tau_2 = R_2 C$

Figure 5.12b shows an active lead-lag filter, whose transfer function is repeated here for convenience

$$F(s) = k_a \frac{1 + s\tau_2}{1 + s\tau_1} \tag{5.32}$$

where

$\tau_1 = R_1 C_1$
$\tau_2 = R_2 C_2$
$k_a = -\frac{C_1}{C_2}$

A "PI" filter is shown in Figure 5.12c. The transfer function is given by

$$F(s) = \frac{1 + s\tau_2}{s\tau_1} \tag{5.33}$$

where

$\tau_1 = R_1 C$
$\tau_2 = R_2 C$

Their Bode plots are shown in Figure 5.13a through c, respectively. High order filters could be used in some applications, but additional filter poles introduce a phase shift. In general, it is not trivial to maintain the stability of high order systems.

The transfer functions of the LFs shown in Figure 5.12 are substituted for $F(s)$ in Equation 5.2 in order to analyze the phase transfer function. We obtain the phase transfer functions as follows: for the passive lead-lag filter

$$H(s) = \frac{k_d k_o (1 + s\tau_2/\tau_1 + \tau_2)}{s^2 + s(1 + k_d k_o \tau_2/\tau_1 + \tau_2) + (k_d k_o/\tau_1 + \tau_2)} = \frac{\omega_n(2\zeta - (\omega_n/k_d k_o))s + \omega_n^2}{s^2 + 2s\zeta\omega_n + \omega_n^2} \tag{5.34}$$

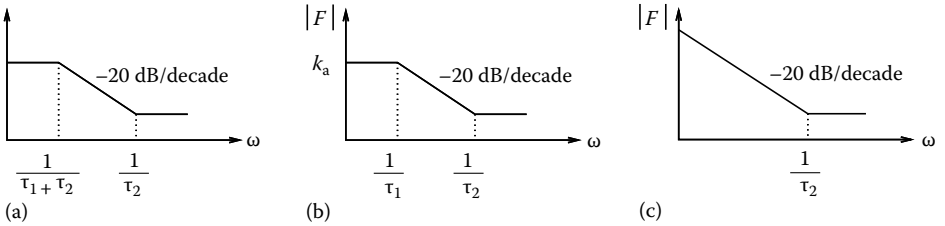


FIGURE 5.13 Bode plots of (a) passive lead-lag filter, (b) active lead-lag filter, and (c) active PI filter.

for the active lead-lag filter

$$H(s) = \frac{k_d k_a k_o (1 + s\tau_2/\tau_1)}{s^2 + s(1 + k_d k_a k_o \tau_2/\tau_1) + (k_d k_a k_o/\tau_1)} = \frac{\omega_n(2\zeta - (\omega_n/k_d k_a k_o))s + \omega_n^2}{s^2 + 2s\zeta\omega_n + \omega_n^2} \tag{5.35}$$

and for the active PI filter

$$H(s) = \frac{k_d k_o (1 + s\tau_2/\tau_1)}{s^2 + s(k_d k_o \tau_2/\tau_1) + (k_d k_o/\tau_1 + \tau_2)} = \frac{2\zeta\omega_n s + \omega_n^2}{s^2 + 2s\zeta\omega_n + \omega_n^2} \tag{5.36}$$

If the condition $k_d k_o \gg \omega_n$ or $k_d k_o k_a \gg \omega_n$ is true, this PLL system is called a "high gain loop." If the reverse is true, the system is a "low gain loop." Most practical PLLs are a high gain loop for good tracking performance. For a high gain loop, Equations 5.34 through 5.36 become approximately

$$H(s) \approx \frac{2\zeta\omega_n s + \omega_n^2}{s^2 + 2s\zeta\omega_n + \omega_n^2} \tag{5.37}$$

Similarly, assuming a high gain loop, the approximate expression of the phase-error transfer function $H_e(s)$ for all three LF types becomes

$$H_e(s) \approx \frac{s^2}{s^2 + 2s\zeta\omega_n + \omega_n^2} \tag{5.38}$$

The magnitude frequency responses of $H(s)$ for a high gain loop with several values of damping factor are plotted in Figure 5.14. It exhibits that the loop performs an LPF on the input phase signal. That is, the second-order PLL is able to track both phase and frequency modulations of the input signal as long as the modulation frequency remains within the frequency band roughly between zero and ω_n.

The transfer function $H(s)$ has a −3 dB frequency, $\omega_{-3\text{ dB}}$, that stands for the close loop bandwidth of the PLL. The relationship between $\omega_{-3\text{ dB}}$ and ω_n is presented here to provide a comparison with a familiar concept of bandwidth.

In a high gain loop case, by setting $|H(j\omega)| = \frac{1}{\sqrt{2}}$ and solving for ω, we can find

$$\omega_{-3\text{ dB}} = \omega_n\left[2\zeta^2 + 1 + \sqrt{(2\zeta^2 + 1)^2}\right]^{\frac{1}{2}} \tag{5.39}$$

The relationship between $\omega_{-3\text{ dB}}$ and ω_n for different damping factors is plotted in Figure 5.15 [4].

The magnitude frequency responses of $H_e(s)$ are plotted in Figure 5.16. A high pass characteristic is observed. It indicates that the second-order PLL tracks the low frequency phase error but cannot track high frequency phase error.

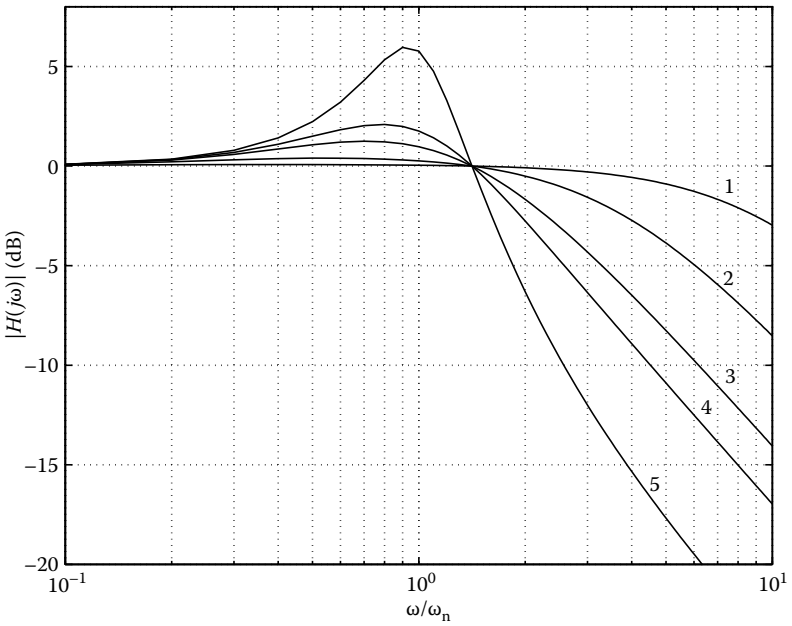


FIGURE 5.14 Frequency responses of the phase transfer function $H(j\omega)$ for different damping factors. Trace1: $\zeta = 5$, Trace2: $\zeta = 2$, Trace3: $\zeta = 1$, Trace4: $\zeta = 0.707$, Trace5: $\zeta = 0.3$.

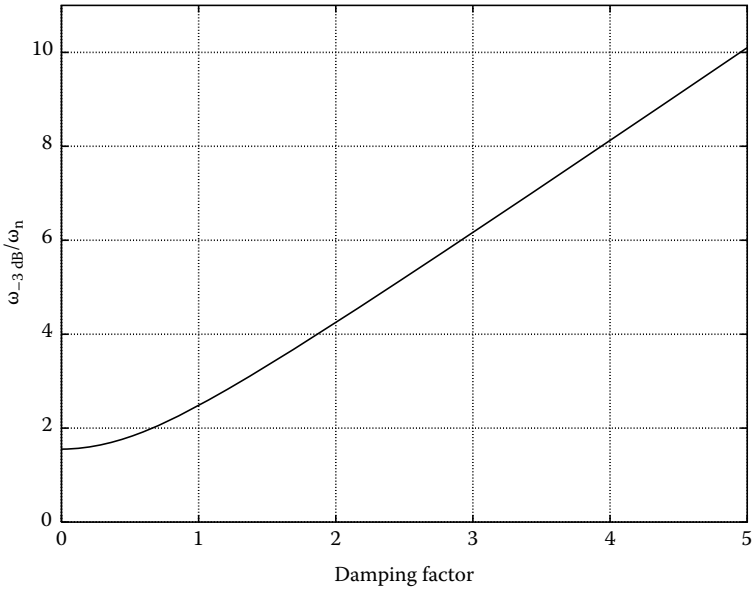


FIGURE 5.15 $\omega_{-3\ dB}$ bandwidth of a second-order loop versus different damping factors.

5.3.3.2 Transformations from *s*-Domain to *z*-Domain

As mentioned in the Section 5.1.3, the function blocks of the ADPLL is implemented by purely digital circuits, and the signals within the loop are digital too. In addition, the SPLL implemented by a microcontroller, microcomputer, or DSP is another type of PLL in discrete-domain (*z*-domain).

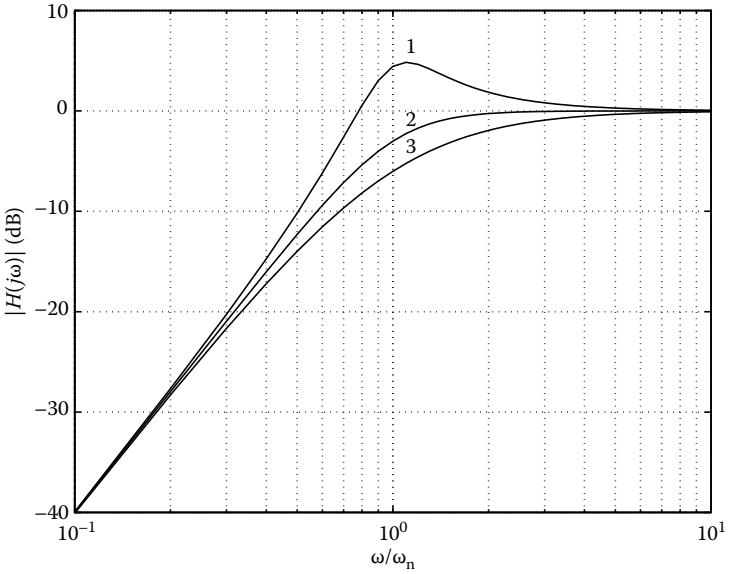


FIGURE 5.16 Frequency responses of the phase-error transfer function $H_e(j\omega)$ for different damping factors. Trace1: $\zeta = 0.3$, Trace2: $\zeta = 0.707$, Trace3: $\zeta = 1$.

Therefore, the analysis and design of a PLL had better be in discrete-domain. The basic types of LF and their features have been described in Section 5.3.3.1. Here, the corresponding discrete-time version of the three basic types of LF will be described after the introduction of transformations from continuous-domain (*s*-domain) to discrete-domain. There are two popular methods to transform a filter from continuous-domain to discrete-domain: backward difference method and bilinear transformation method. Figure 5.17 shows the principle of the backward difference method that we approximate the area under each segment of continuous curve by a rectangular area. Referring to Figure 5.17, the backward difference method means to approximate the integration areas of $\int_{(k-1)}^{kT} y(t)\mathrm{d}t$ by $y(kT)T$. Based on the backward difference method, the *z*-domain equivalent transfer function $H(z)$ of an *s*-domain transfer function $H(s)$ is simple and obtained by the substitution

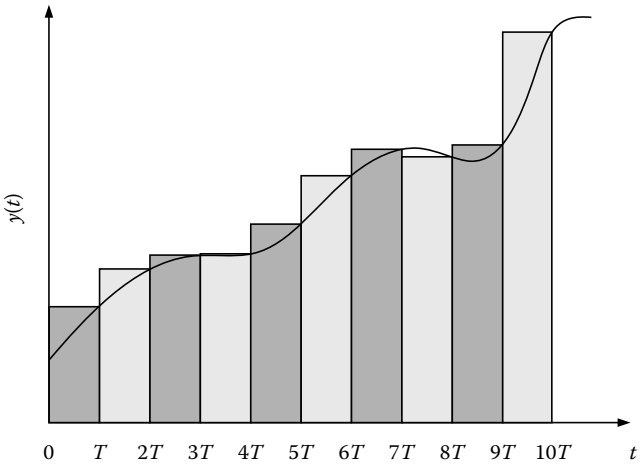


FIGURE 5.17 Backward difference method using a rectangular area approximation.

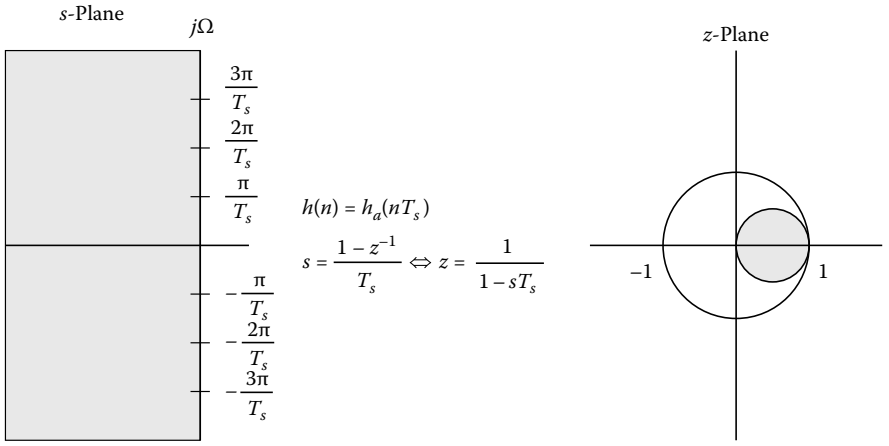


FIGURE 5.18 Mapping of the left half of the s-plane into the z-plane by the backward difference method.

$$H(z) = H(s)\big|_{s=\frac{1-z^{-1}}{T_s}} \tag{5.40}$$

One of the advantages of the backward difference method is that it will produce a stable discrete-time filter for a stable continuous-time filter. Figure 5.18 shows the mapping of the left half of the s-plane into the z-plane by the backward difference method. However, there is considerable distortion in the transient and frequency response characteristics of the discrete-time filter obtained in this method since the stable region is mapped into only a circle within the unit circle of z-plane.

From calculus and Figure 5.17, a good approximation is obtained only if the continuous-time signal changes very slowly over the sampling interval T_s. In other words, the signal bandwidth has to be much smaller than the sampling rate since the mapping from s-domain to z-domain should become distorted while the sampling period is too long. To reduce the distortion, it is desired to use a faster sampling frequency, that is, a smaller sampling period.

Figure 5.19 shows the principle of the bilinear transformation method that we approximate the area under each segment of continuous curve by a trapezoidal area. Therefore, the bilinear transformation method is also called the trapezoidal integration method to approximate the integration areas $\int_{(k-1)}^{kT} y(t)\mathrm{d}t$

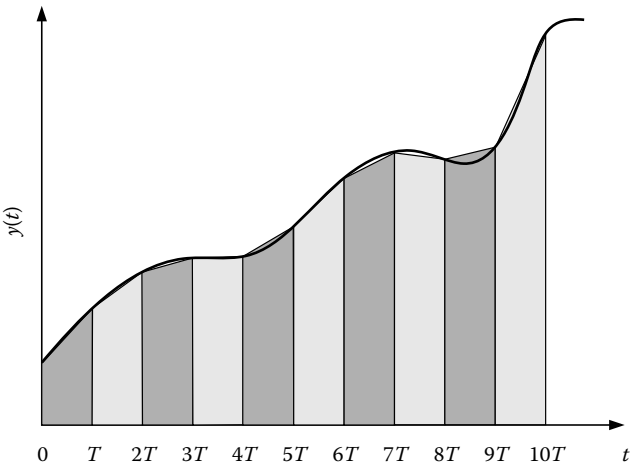


FIGURE 5.19 Bilinear transformation method using trapezoidal area approximation.

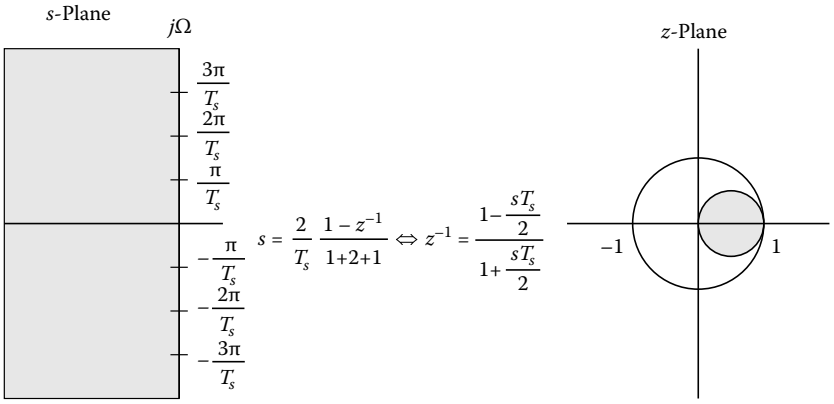


FIGURE 5.20 Mapping of the left half of the *s*-plane into the *z*-plane by the bilinear transformation method.

by $\frac{1}{2}[y(kT) + y((k-1)T)]T$. Thus, the *z*-domain equivalent transfer function $H(z)$ of a continuous-time filter $H(s)$ is obtained by

$$H(z) = H(s)\big|_{s=\frac{2}{T_s}\frac{1-z^{-1}}{1+z^{-1}}} \tag{5.41}$$

By means of bilinear transformation method, the entire left half of the *s*-plane is mapped into the unit circle with center at the origin of the *z*-plane as shown in Figure 5.20. Hence, the bilinear transformation method produces a stable discrete-time filter for a stable continuous-time filter. Furthermore, there is no frequency folding by means of the bilinear transformation method since it maps the entire *jw* axis of the *s*-plane into one complete revolution of the unit circle in the *z*-plane.

5.3.3.3 Discrete-Time LFs

As mentioned in the Section 5.3.3.1, there are three typical LFs for a phase-locked filter. Using backward difference transformation, the discrete-time transfer functions of passive lead-lag filter, active lead-lag filter, and active PI filter can be obtained as follows:

Passive lead-lag filter:

$$F_{\text{back,PLL}}(z) = \frac{T_s + \tau_2 - \tau_2 z^{-1}}{T_s + \tau_1 + \tau_2 - (\tau_1 + \tau_2)z^{-1}} \tag{5.42}$$

Active lead-lag filter:

$$F_{\text{back,ALL}}(z) = k_a \frac{T_s + \tau_2 - \tau_2 z^{-1}}{T_s + \tau_1 - \tau_1 z^{-1}} \tag{5.43}$$

Active PI filter:

$$F_{\text{back,PI}}(z) = \frac{T_s + \tau_2 - \tau_2 z^{-1}}{\tau_1(1 - z^{-1})} \tag{5.44}$$

On the other hand, the discrete-time transfer functions of passive lead-lag filter, active lead-lag filter, and active PI filter can be written, using the bilinear transformation, as follows:

Passive lead-lag filter:

$$F_{\text{bilinear,PLL}}(z) = \frac{T_s + 2\tau_2 + (T_s - 2\tau_2)z^{-1}}{T_s + 2\tau_1 + 2\tau_2 + (T_s - 2\tau_1 - 2\tau_2)z^{-1}} \tag{5.45}$$

Active lead-lag filter:

$$F_{\text{bilinear,ALL}}(z) = k_a \frac{T_s + 2\tau_2 + (T_s - 2\tau_2)z^{-1}}{T_s + 2\tau_1 + (T_s - 2\tau_1)z^{-1}} \tag{5.46}$$

Active PI filter:

$$F_{\text{bilinear}}(z) = \frac{T_s + 2\tau_2 + (T_s - 2\tau_2)z^{-1}}{2\tau_1(1 - z^{-1})} \tag{5.47}$$

From the viewpoint of implementation, all the discrete-time LFs have the transfer function format of a first-order infinite impulse response filter:

$$F_{\text{LF}}(z) = \frac{b_0 + b_1 z^{-1}}{1 - a_1 z^{-1}} \tag{5.48}$$

The differences in hardware requirements are small, but the system characteristics and performance are dramatic. Using an approach similar to the Weiner filter theory, Jaffe and Rechtin [30] investigated the optimal LFs for PLLs with different inputs. For a frequency step input, the form of active PI filter is shown to be optimal.

5.3.4 Charge-Pump PLL

A charge-pump PLL usually consists of four major blocks as shown in Figure 5.21. The PD is a purely PFD. The charge-pump circuit converts the digital signals UP, DN, and null (neither up nor down) generated by the PD into a corresponding charge-pump current I_p, $-I_p$, and zero. The LF is usually a passive RC circuit converting the charge-pump current into an analog voltage to control VCO. The purpose of the "charge-pump" is to convert the logic state of the phase-frequency detector output into an analog signal suitable for controlling the VCO. The schematic of the charge-pump circuit and the LF is shown in Figure 5.22. The linear model shown in Figure 5.3 can be employed to describe a charge-pump PLL. k_d is the equivalent gain of a charge-pump circuit. If the loop bandwidth is much smaller than the input frequency, the detailed behavior within a single cycle can be ignored. Then the state of a PLL can be assumed to be only changed by a small amount during each input cycle. Actually the "average" behavior over many cycles is what we are interested in. The average current charging the capacitor is given by

$$\begin{aligned} I_{\text{avg}} &= \frac{Q}{T} = \frac{I\Delta t}{T} \\ &= \frac{I\left(\frac{\phi_e}{2\pi}\right)T}{T} \\ &= \frac{I\phi_e}{2\pi} \end{aligned} \tag{5.49}$$

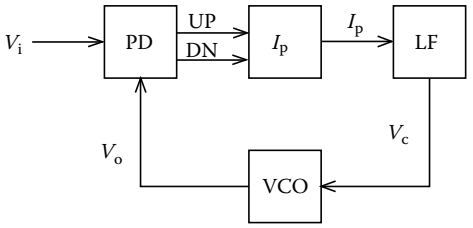


FIGURE 5.21 Charge-pump PLL diagram.

And the average k_d is

$$k_d \triangleq \frac{I_{\text{avg}}}{\phi_e} = \frac{I_{\text{avg}}}{2\pi} \tag{5.50}$$

The charge-pump current is transferred to the control voltage of the following VCO by the LF consisted of a resistor and a capacitor as shown in Figure 5.22. The impedance (transfer function) of the RC LF is given by

$$F(s) = R + \frac{1}{C_p s} = \frac{1 + RC_p s}{C_p s} = K_p + \frac{K_I}{s} \tag{5.51}$$

which has the format of an active PI filter. Therefore, the closed-loop transfer function can be obtained as

$$H(s) \triangleq \frac{\phi_{out}}{\phi_{in}} = \frac{k_d \cdot F(s) \cdot \frac{k_o}{s}}{1 + \frac{k_d k_o}{s}} = \frac{\frac{I_{avg}}{2\pi C_p}(RC_p s + 1)k_o}{s^2 + \frac{I_{avg}}{2\pi} k_o R s + \frac{I_{avg}}{2\pi C_p} k_o} \tag{5.52}$$

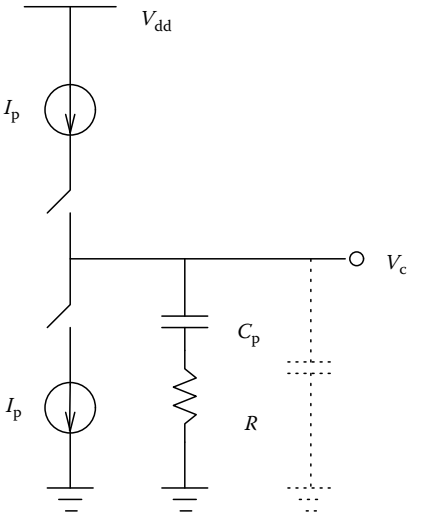


FIGURE 5.22 The schematic of LF.

Generally, a second-order system is characterized by the natural frequency $f_n = \frac{\omega n}{2\pi}$ and the damping factor ζ, and they can be expressed as follows:

$$\omega_n = \sqrt{\frac{I_{avg}}{2\pi C_p} k_o} \text{ rad/s}$$
$$\zeta = \frac{RC_p}{2}\omega_n \tag{5.53}$$

For the stability consideration, there is a limitation of a normalized natural frequency F_N [15],

$$F_N \triangleq \frac{f_n}{f_i} < \frac{\sqrt{1 + \zeta^2} - \zeta}{\pi} \tag{5.54}$$

In the single-ended charge pump, the resistor added in series with the capacitor shown in Figure 5.22 may introduce ''ripple'' in the control voltage V_c even when the loop is locked [16]. The ripple control voltage modulates the VCO frequency and results in phase noise. This effect is especially undesired in frequency synthesizers. In order to suppress the ripple, a second-order LF, as shown in Figure 5.22 with a shunt capacitor in dotted line, is used. This configuration introduces a third pole in the PLL. Stability issues must be taken care of furthermore. Gardner provides criteria for the stability of the third-order PLL [16].

An important property of any PLLs is the static phase error that arises from a frequency offset $\Delta\omega$ between the input signal and the free-running frequency of the VCO. According to the analysis in Ref. [16], the static phase error is

$$\theta_v = \frac{2\pi\Delta\omega}{k_o I_p F(0)} \text{ rad} \tag{5.55}$$

To eliminate the static phase error in conventional PLLs, an active LF with a high DC gain ($F(0)$ is large) is preferred. Nevertheless, the charge-pump PLL allows zero static phase error without the need of

a large DC gain of the LF. This effect arises from the input open circuit during the "null" state (charge-pump current is zero). Real circuits will impose some resistive loading R_s in parallel to the LF. Therefore, the static phase error, from Equation 5.55 will be

$$\theta_v = \frac{2\pi\Delta\omega}{k_o I_p R_s} \text{ rad} \tag{5.56}$$

The shunt resistive loading most likely comes from the input of a VCO control terminal. Compared with the static phase error of a conventional PLL as expressed in Equation 5.11, the same performance can be obtained from a charge-pump PLL without a high DC-gain LF [30].

5.3.5 PLL Design Considerations

5.3.5.1 Typical Procedures of PLL Design

A PLL design usually starts with specifying the key parameters such as natural frequency ω_n, lock-in range $\Delta\omega_L$, damping factor ζ, and the frequency control range which majorly depend on applications. Typical design procedures are described as follows:

Step 1. Specify the damping factor ζ. The damping factor determines the relative stability of a PLL. ζ should be considered as a critical parameter to achieve fast response, small overshoot, and minimum noise bandwidth B_L. If ζ is very small, large overshoot occurs and the overshoot causes phase jitter [19]. If ζ is too large, the response becomes sluggish.

Step 2. Specify the lock-in range $\Delta\omega_L$ or the noise bandwidth B_L. As shown in Equations 5.53 and 5.50, the natural frequency ω_n depends on $\Delta\omega_L$ and ζ (or B_L and ζ). If the noise is not the key issue of the PLL, we may ignore the noise bandwidth and specify the lock-in range. If the noise is concerned, we should specify B_L first, and keep the lock-in range of PLL.

Step 3. Calculate the ω_n according to step 2. If the lock-in range has been specified, Equation 5.53 indicates

$$\omega_n = \frac{\Delta\omega_L}{2\zeta} \tag{5.57}$$

If the noise bandwidth has been specified, Equation 5.50 indicates the natural frequency as

$$\omega_n = \frac{2B_L}{\zeta + \frac{1}{4\zeta}} \tag{5.58}$$

Step 4. Determine the VCO gain factor k_o and the PD gain k_d. k_o and k_d are both characterized by circuit architectures. They must achieve the requirement of the lock-in range specified in step 2. For example, if k_o or k_d is too small, the PLL will fail to achieve the desired lock-in range.

Step 5. Choose the LF. Different types of the LF are available as shown in Figure 5.12. According to Equation 5.9 through 5.11, ω_n and ζ specified above are used to derive the time constants of the LF.

5.3.5.2 PLL Bandwidth Control

As illustrated in previous sections, the PLL noise performance is contrary to the dynamic performances of a PLL such as lock-in range and acquisition speed, which are generally proportional to the PLL bandwidth or the nature frequency ω_n. For example, Equation 5.52 shows that it is desired to have large k_d or $F(\Delta\omega_L)$ to achieve wide lock-in range and fast acquisition. In contrast, Equations 5.59 and 5.60 show that it is desired to have a small nature frequency ω_n in order to reduce the output jitter. Besides, variations in process, voltage, and temperature can lead to uncertainties in loop parameters of a

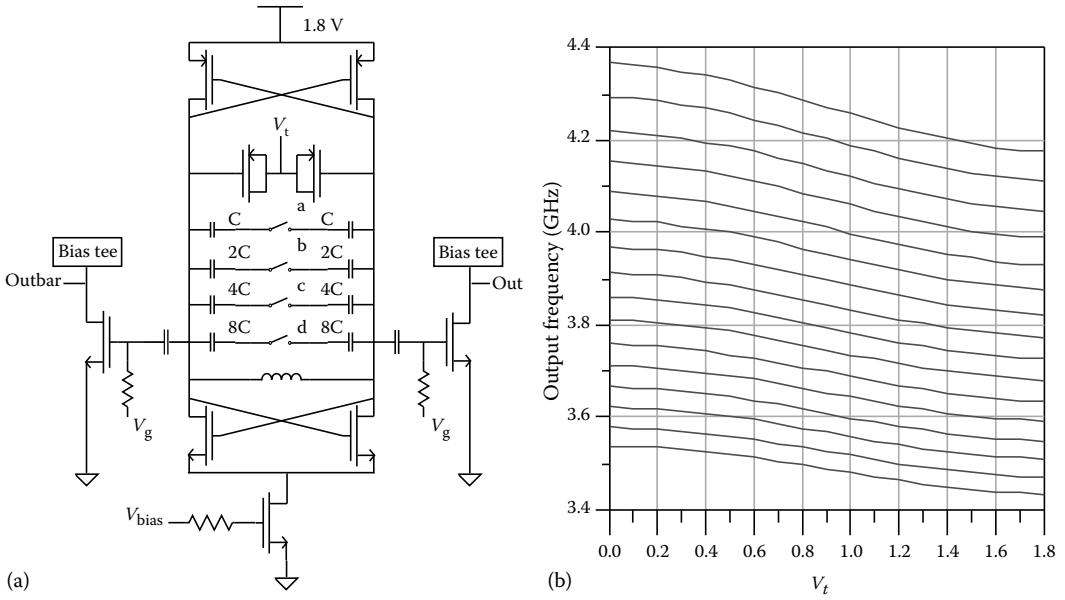


FIGURE 5.23 (a) Wide-range LC-tank VCO with capacitor array and (b) conversion gain k_o.

fixed-bandwidth PLL, such as the conversion gain k_o of VCO, the time-constant ratio τ_2/τ_1 of the LF, and the charge-pump current I_p. For example, variations in process, voltage, and temperature will cause the frequency of a ring oscillator to vary by a factor of 2–3 between its slowest and fastest conditions experientially. Figure 5.23 shows a wide-range LC-tank VCO with capacitor array to achieve low conversion gain k_o for superior phase noise performance. It can be seen that the VCO conversion gain k_o slightly varies from segment to segment of frequencies. Moreover, the curve of conversion gain k_o is nonlinear in any frequency segment. The design result shows a conversion gain range of 27–50 MHz/V. This means that the PLL bandwidth will vary by a factor of 2 due to the variation of VCO conversion gain k_o. Following the design procedures mentioned above, a conservative operating point is usually chosen to guarantee the stability for all conditions. Unfortunately, such a conservative design only achieves a suboptimal performance in most cases.

For achieving fast acquisition as well as a superior noise performance, many schemes of PLL bandwidth control are proposed in the literature [31–37]. Figure 5.24 shows the concept of PLL bandwidth control based on some monitoring techniques of the phase error. In Ref. [31], based on some modifications of the Kalman filtering formulation, an approach is developed to derive the optimal loop gain sequence of dual-loop DPLL, which is independent of measured noise statistics. In order to achieve low noise PLL as

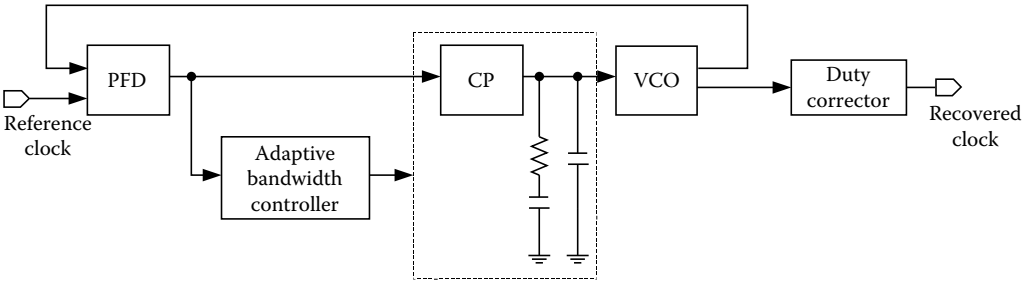


FIGURE 5.24 The concept of an adaptive-bandwidth PLL.

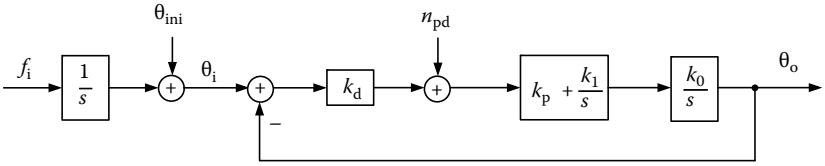


FIGURE 5.25 Linear model of a charge-pump PLL.

well as fast acquistion, Joonsuk Lee and Beomsup Kim [32] realize an analog bandwidth controller to adaptively control the current of charge-pump, that is the PLL bandwidth. Figure 5.25 shows the linear model of charge-pump PLL using a proportional–integral (PI) filter as shown in Figure 5.21. According to Equations 5.51 through 5.53, the damping factor ζ, that is, relative stability of such an adaptive-bandwidth PLL will vary during the lock-in process since the nature frequency ω_n is adaptively adjusted.

In order to develop a PLL and delay-locked loop (DLL) with adaptive-bandwidth to enable optimal performance over a wide frequency range and across process, voltage, and temperature variations, a discrete-time, open-loop dynamic model of the PLL/DLL is proposed to characterize the change in output variables in response to the sampled error, and to express the adaptive-bandwidth criteria in terms of the open-loop gains [37]. Furthermore, the scaling equations for the charge-pump current and the filter resistor are derived to achieve adaptive-bandwidth charge-pump PLL/DLLs with a constant damping factor ζ. It is regrettable that such scaling scheme is a challenge in the analog realization. In contrast, it is easier to maintain a constant damping factor ζ using a digital PI LF to achieve the adaptive bandwidth control. In Ref. [11], a modified structure of digital PI LF is proposed and shown in Figure 5.26. Based on such a modified structure, it is easy to derive the equivalent phase-transfer function in continuous-time domain that is represented in terms of the parameters of loop components as

$$\begin{aligned} H_\theta(s) &= \frac{\Theta_o(s)}{\Theta_i(s)} \\ &= \frac{k_d k_o K_P s + k_d k_o K_P K_I}{s^2 + k_d k_o K_P s + k_d k_o K_P K_I} \\ &= \frac{2\zeta\omega_n s + \omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \end{aligned} \tag{5.59}$$

where

$$\begin{aligned} \omega_n &= \sqrt{k_d k_o K_P K_I}, \\ \zeta &= 0.5\sqrt{k_d k_o K_P / K_I} \end{aligned} \tag{5.60}$$

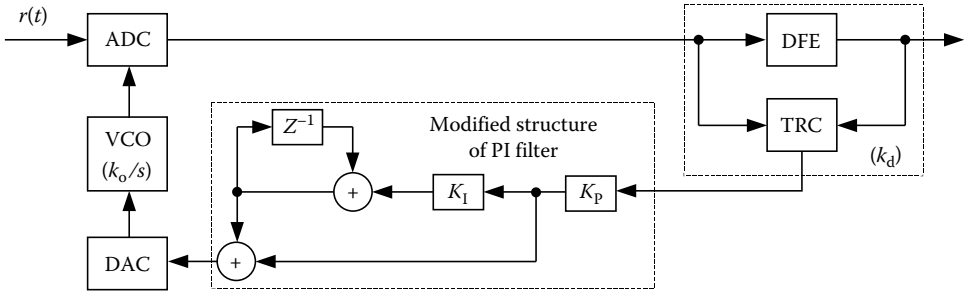


FIGURE 5.26 Timing recovery using a modified structure of digital PI LF.

denote the nature frequency and the damping factor, respectively. Referring to Equation 5.60, if a scaling factor of K_P and K_I is induced into the modified PI LF structure for adaptively adjusting the PLL bandwidth or the nature frequency ω_n, we can mantain a constant damping factor ζ while the PLL bandwidth is proportional to the scaling factor induced.

5.4 PLL Applications

5.4.1 Clock and Data Recovery

In data transmission systems such as optical communications, telecommunications, disk drive systems, and local networks, data are transmitted on baseband or passband. In most of these applications, only data signals are transmitted by transmitter, but clock signals are not transmitted in order to save hardware cost. Therefore, the receiver should have some schemes to extract the clock information from the received data stream and to regenerate transmitted data using the recovered clock. This scheme is called timing recovery or clock recovery.

To recover the data correctly, the receiver must generate a synchronous clock from the input data stream, the recovered clock must synchronize with the bit rate (the baud of data). The PLL can be used to recover the clock from the data stream, but there are some special design considerations. For example, because of the random nature of data, the choice of PFDs is restricted. In particular, three-state PD is not proper, because when there are no transitions in the data stream, the PD interprets that the VCO frequency is higher than the data frequency and remains its output in the ''down'' state, which makes the PLL lose lock as shown in Figure 5.27. Thus, the choice of PFD for random binary data requires a careful examination over whether data transitions are absent. One useful method is the rotational frequency detector described in Ref. [9]. The random data also causes the PLL to introduce undesired phase variation in the recovered clock, it is called timing jitter and this is an important issue of the clock recovery.

5.4.1.1 Data Format

Binary data are usually transmitted in an NRZ format as shown in Figure 5.28a because of the consideration of bandwidth efficiency. In NRZ format, each bit has a duration of T_B (bit period). The signal does not go to zero between adjacent pulses representing 1's. It can be shown in Ref. [24] that the corresponding spectrum has no line component at $f_B = \frac{1}{T_B}$, most of the spectrum of this signal lines below $\frac{f_B}{2}$. The term ''NRZ'' distinguishes itself from another data type called ''return-to-zero'' (RZ) as shown in Figure 5.28b, where the signal goes to zero between consecutive bits. Therefore, the spectrum of RZ data have a frequency component at f_B. For a given bit rate, RZ data need wider transmitting bandwidth; therefore, NRZ data are preferable when channel or circuit bandwidth is a concern.

Due to the lack of a spectral component at the bit rate of NRZ format, a clock recovery circuit may lock to spurious signals or fail to lock at all. Thus, a nonlinear process for the NRZ data is essential to create a frequency component at the baud rate.

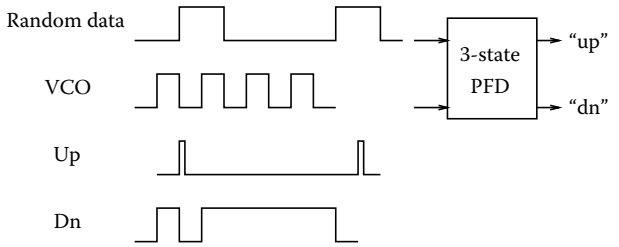


FIGURE 5.27 Response of a three-state PD to random data.

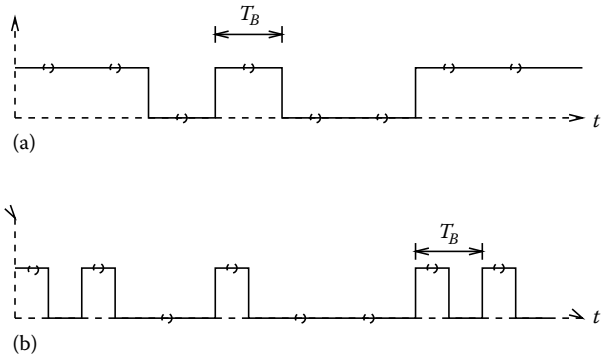


FIGURE 5.28 (a) NRZ data and (b) RZ data.

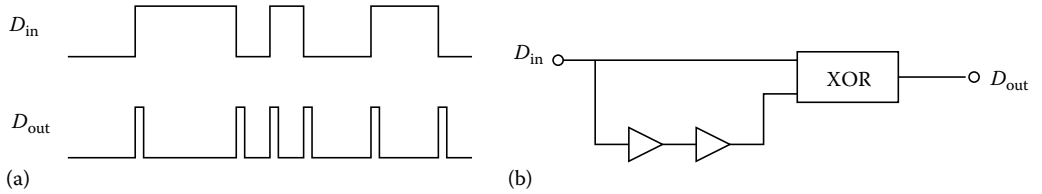


FIGURE 5.29 Edge detection of NRZ data.

5.4.1.2 Data Conversion

One way to recover the clock signal from the NRZ data is to convert it to an RZ-like data that have a frequency component at the bit rate, and then recover clock from data using a PLL. Transition detection is one of the methods to convert NRZ data to RZ-like data. As illustrated in Figure 5.29a, the edge detection requires a mechanism to sense both positive and negative data transitions. In Figure 5.29b, NRZ data are delayed and compared with itself by an exclusive-OR gate; therefore, the transition edges are detected. In Figure 5.30, the NRZ data V_i is first differentiated to generate pulses corresponding to each transition. These pulses are made to be all positive by squaring the differentiated signal v_i. The result is that the signal V_i' looks just like RZ data, where pulses are spaced at an interval of T_B.

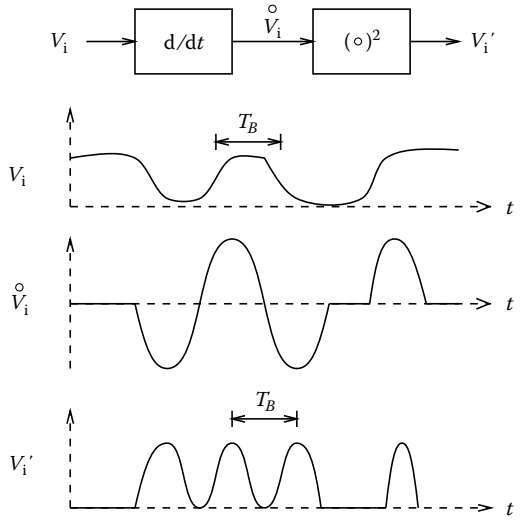


FIGURE 5.30 Converting NRZ to RZ-like signal.

5.4.1.3 Clock Recovery Architecture

Based on different PLL topologies, there are several clock recovery approaches. Here, the early–late and the edge-detector-based methods will be described.

Figure 5.31 shows the block diagram of the early–late method. The waveforms for the case in which the input lags the VCO output are shown in Figure 5.32, where the early integrator integrates the input signal for the early-half period of the clock signal, and holds it for the remainder of the clock signal. On the other

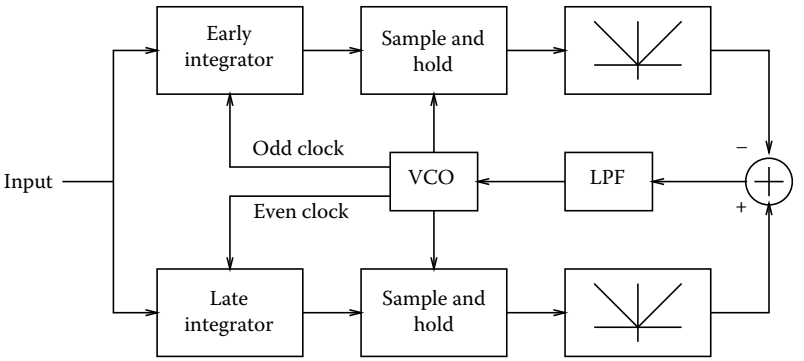


FIGURE 5.31 Early–late block diagram.

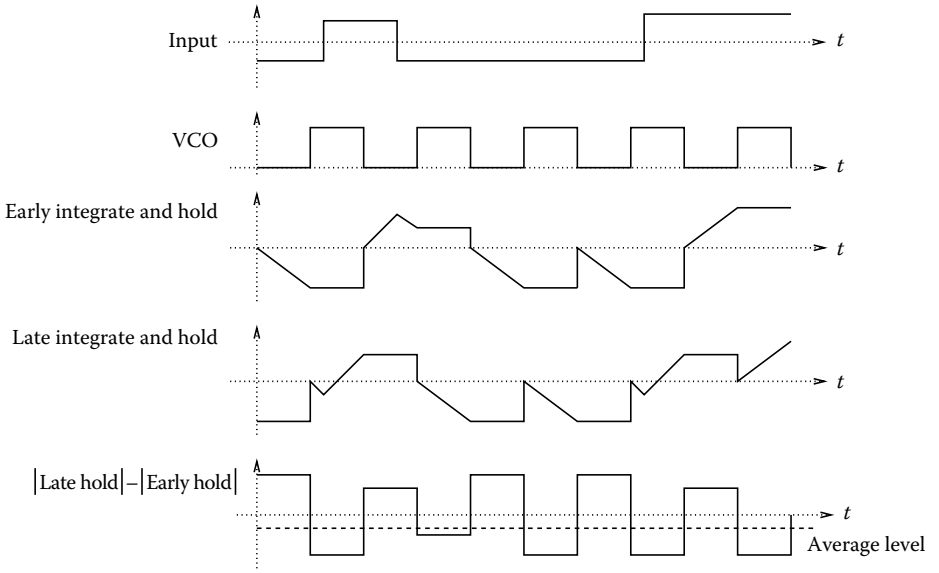


FIGURE 5.32 Clock waveforms for early–late architecture.

hand, the late integrator integrates the input signal for the late-half period of the clock signal and holds it for the next early-half period. The average difference between the absolute values of the late hold and the early hold voltage generated from an LPF gives the control signal to adjust the frequency of the VCO. As mentioned above, this method is popular for rectangular pulses. However, there are some drawbacks in this method. As this method relies on the shape of pulses, a static phase error can be introduced if the pulse shape is not symmetric. In high-speed applications, this approach requires a fast settling integrator that limits the operating speed of the clock recovery circuit and the acquisition time cannot be easily controlled.

The most widely used technique for clock recovery in high performance, wide-band data transmission applications is the edge-detection-based method. The edge-detection method is used to convert data format such that the PLL can lock the correct baud frequency. More details have been described in Section 5.4.1.2. There are many variations of this method depending on the exact implementation of each PLL loop component. The ''quadricorrelator'' introduced by Richman [7] and modified

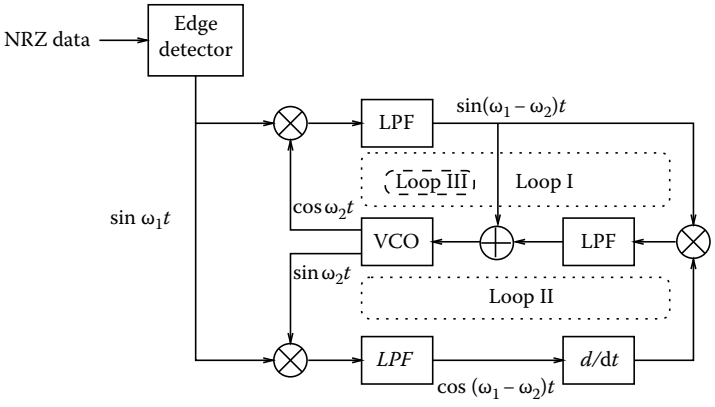


FIGURE 5.33 Quadricorrelator.

by Bellisio [25] is a frequency-difference discriminator and has been implemented in a clock recovery architecture. Figure 5.33 [26] is a phase-frequency locked loop using edge-detection method and quadricorrelator to recover timing information from NRZ data. As shown in Figure 5.33, the quadricorrelator follows the edge-detector with a combination of three loops sharing the same VCO. Loops I and II form a frequency-locked loop that contains the quadricorrelator for frequency detection. Loop III is a typical PLL for phase alignment. Since the phase- and frequency-locked loops share the same VCO, the interaction between two loops is a very important issue. As described in Ref. [26], when $\omega_1 \approx \omega_2$, the DC feedback signal produced by loops I and II approaches zero and loop III dominates the loop performance. A composite frequency- and PLL is a good method to achieve fast acquisition and a narrow PLL loop bandwidth to minimize the VCO drift. Nevertheless, because the wide band frequency-locked loop can response to noise and spurious components, it is essential to disable frequency-locked loop when the frequency error gets into the lock-in range of the PLL to minimize the interaction. More clock recovery architectures are described in Refs. [19,21,23,27–29].

5.4.2 Delay-Locked Loop

Two major elements for adjusting the timing are VCO and voltage-controlled delay line (VCDL). Figure 5.34 shows a typical DLL [12,13] that replaces the VCO of a PLL with a VCDL. The input signal is delayed by an integer multiple of the signal period because the phase error is zero when the phase difference between V_{in} and V_o approaches multiple of the signal periods. The VCDL usually consists a number of cascaded gain stages with variable delay. Delay lines, unlike ring oscillators, cannot generate a signal; therefore, it is difficult to make frequency multiplication in a DLL.

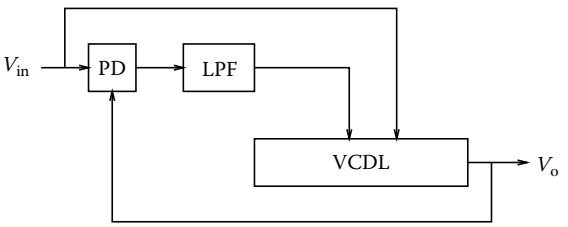


FIGURE 5.34 DLL block diagram.

In a VCO, the output "frequency" is proportional to the input control voltage. The phase transfer function contains a pole, which is $H(s) = \frac{k_o}{s}$ (k_o is the VCO gain). In a VCDL, the output "phase" is proportional to the control voltage, and the phase transfer function is $H(s) = k_{VCDL}$. So the DLL can be easily stabilized with a simple first-order LF. Consequently, DLLs have much more relaxed trade-offs among gain, bandwidth, and stability. This is one of the two important

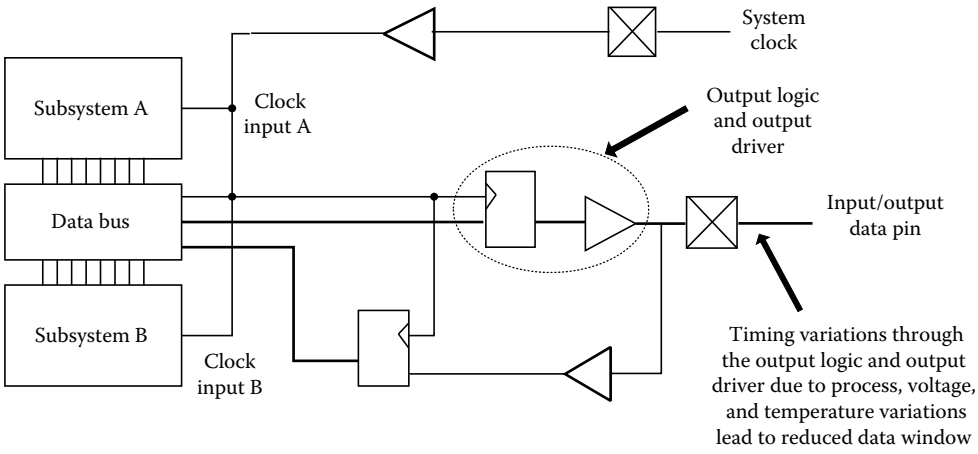

System clock
Subsystem A
Clock input A
Output logic and output driver
Data bus
Input/output data pin
Subsystem B
Clock input B
Timing variations through the output logic and output driver due to process, voltage, and temperature variations lead to reduced data window

FIGURE 5.35 Modern digital systems use synchronous communication to achieve high-speed signaling to and from the bus between the subsystems.

advantages over PLLs. Another advantage is that delay lines typically introduce much less jitter than VCO [14]. Because a delay chain is not configured as a ring-oscillator, there is no jitter accumulation since the noise does not contribute to the starting point of the next clock cycle.

A typical application of DLL is to synchronize the clock edges of subsystems within a digital system to access the bus between subsystems. Figure 5.35 shows modern digital systems that use synchronous communication to achieve high-speed signaling to and from the bus between the subsystems. Subsystems that communicate synchronously use a clock signal as a timing reference so that data can be transmitted and received with a known relationship to this reference. A difficulty in maintaining this relationship is that process, voltage, and temperature variations can alter the timing relationship between the clock and data signals of subsystems, resulting in reduced timing margins. Figure 5.36 shows that on the left side the data valid window (the time over which data can be sampled reliably by the receiver) can be large at low signaling speeds [38]. Even in the presence of a substantial shift in the data valid window across operational extremes, the resulting data valid window can still be large enough to transmit and receive the data reliably. Unfortunately, the variations in process, voltage, and temperature can result in the loss of

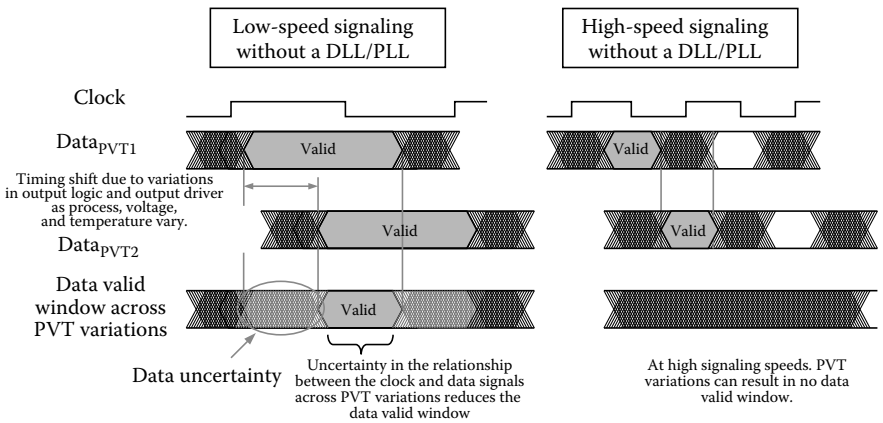

Low-speed signaling without a DLL/PLL
High-speed signaling without a DLL/PLL
Clock
Data$_{PVT1}$
Valid
Valid
Timing shift due to variations in output logic and output driver as process, voltage, and temperature vary.
Data$_{PVT2}$
Valid
Valid
Data valid window across PVT variations
Valid
Data uncertainty
Uncertainty in the relationship between the clock and data signals across PVT variations reduces the data valid window
At high signaling speeds. PVT variations can result in no data valid window.

FIGURE 5.36 The timing relationships between the clock and data signals of subsystems in a conventional digital system.

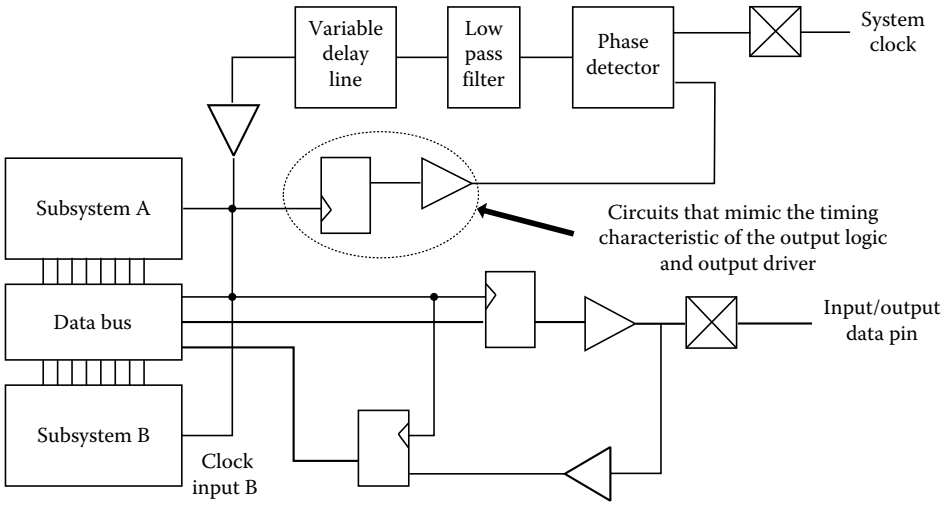


FIGURE 5.37 DLL-on-chip to maintain the timing relationship between a clock signal and an output data signal.

the data valid window when the signal speed is increased as also shown on the right-hand side of Figure 5.36. This problem gets worse as signaling speeds increase, limiting the ability of subsystems to communicate data at higher speeds.

The function of DLLs and PLLs to synchronize a signal with a reference or an input signal in frequency as well as in phase can be used to maintain such a fixed timing relationship between signals of subsystems. Figure 5.37 shows how a DLL is used to maintain the timing relationship between a clock signal and an output data signal. The PD detects phase differences between the clock and output data and sends control information through an LPF to a variable delay line that adjusts the timing of the internal clock to maintain the desired timing relationship. The PD must account for the timing characteristics of the output logic and output driver. This is important since it estimates the phase differences between the clock and the data driven by the output driver, where the timing relationships of subsystems are changed over time due to the process, voltage, and temperature variations. Maintaining the timing relationships between the clock and output data with DLLs and PLLs results in improved timing margins as shown in Figure 5.38. Then, the important limitation to increasing signaling speeds is addressed.

5.4.3 Frequency Synthesizer

A frequency synthesizer generates any of a number of frequencies by locking a VCO to an accurate frequency source such as a crystal oscillator. For example, RF systems usually require a high-frequency local oscillator whose frequency can be changed in small and precise steps. The ability of multiplying a reference frequency makes PLLs attractive for synthesizing frequencies.

The basic configuration used for frequency synthesis is shown in Figure 5.39a. The system is capable of generating the frequency at an integer multiple of the reference frequency. A quartz crystal is usually used as the reference clock source because of its low jitter characteristic. Due to the limited speed of CMOS device, it is difficult to generate frequency directly in the range of GHz or more. To generate higher frequencies, prescalers are used, which are implemented with other IC technologies such as ECL. Figure 5.39b shows a synthesizer structure using a prescaler V, where the output frequency becomes

$$f_{\text{out}} = \frac{NVf_{\text{i}}}{M} \tag{5.61}$$

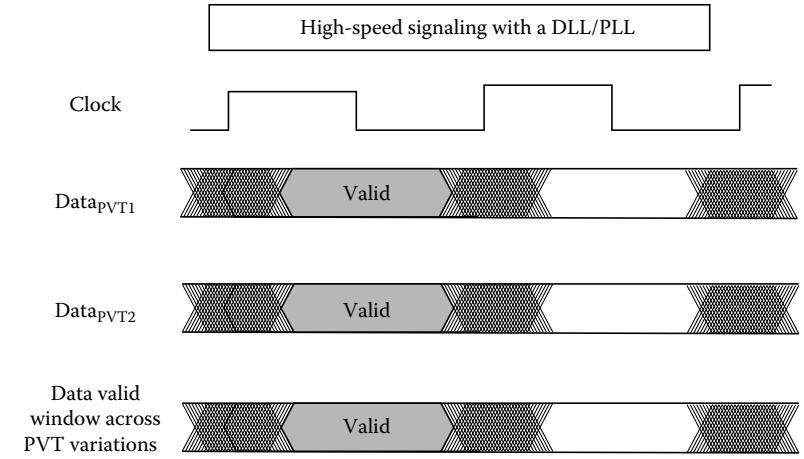


FIGURE 5.38 Timing relationships between subsystems when a DLL is employed for synchronizing the bus access.

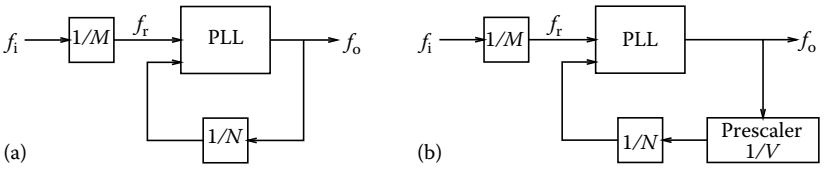


FIGURE 5.39 Frequency-synthesizer block diagrams: (a) basic frequency-synthesizer system and (b) system extends the upper frequency range by using an additional high-speed prescaler.

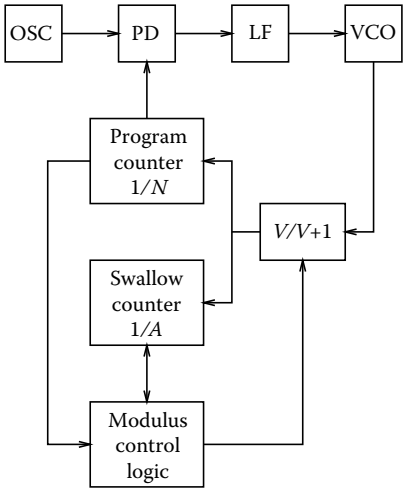


FIGURE 5.40 The block diagram of dual-modulus frequency synthesizer.

Because the scaling factor V is obviously much greater than one, it is no longer possible to generate any desired integer multiple of the reference frequency. This drawback can be circumvented by using a so-called dual-modulus prescaler as shown in Figure 5.40. A dual-modulus prescaler is a divider whose division can be switched from one value to the other by a control signal. The following shows that the dual-modulus prescaler makes it possible to generate a number of output frequencies that are spaced only by one reference frequency. The VCO output is divided by $V/V+1$ dual-modulus prescaler. The output of the prescaler is fed into a "program counter" $1/N$ and a "swallow counter" $1/A$. The dual-modulus prescaler is set to divide by $V+1$ initially. After "A" pulses out of the prescaler, the swallow counter is full and changes the prescaler modulus to V. After additional "N–A" pulses out of the prescaler, the program counter changes the prescaler modulus back to $V+1$ and restarts the swallow counter. Then the cycle

is repeated. In this way, the VCO frequency is equal to $(V+1)A + V(N-A) = VN + A$ times of the reference frequency. Note that N must be larger than A. If this is not the case, the program counter would be full earlier than the $1/A$ and both counters would be reset. Therefore, the dual-modulus prescaler would never be switched from $V+1$ to V. For example, if $V=64$, then A must be in the range of 0–63 such that $N_{\min}=64$. The smallest realizable division ratio is

$$(N_{\text{tot}})_{\min} = N_{\min} V = 4096 \tag{5.62}$$

The synthesizer of Figure 5.40 is able to generate all integer multiple of the reference frequency starting from $N_{\text{tot}}=4096$. For extending the upper frequency range of frequency synthesizers but still allowing the synthesis of lower frequency, the four-modulus prescaler is a solution [1].

Based on the above discussions, the synthesized frequency is an integer multiple of a reference frequency. In RF applications, the reference frequency is usually larger than the channel spacing for loop dynamic performance considerations, in which the wider loop bandwidth for a given channel spacing allows faster settling time and reduces the phase jitter requirements to be imposed on the VCO. Therefore a ''fractional'' scaling factor is needed. Fractional division ratios of any complexity can be realized. For example, a ratio 3.7 is obtained if a counter is forced to divide by 4 in seven cycles of each group of 10 cycles and by 3 in the remaining three cycles. On the average, this counter divides the input frequency by 3.7 effectively.

Bibliography

1. R. E. Best, *Phase-Locked Loops Theory, Design, Applications*, McGraw-Hill, New York, 1984.
2. T. G. Donald and J. D. Gallia, *Digital Phase-Locked Loop Design Using SN54/74LS297*, Application Note AN 3216, Texas Instruments Inc., Dallas, TX.
3. W. Rosink, All-digital phase-locked loops using the 74HC = HCT297, *Electron. Components and Appl.*, 9, 66–89, 1989.
4. F. M. Gardner, *Phaselock Techniques*, 2nd ed., John Wiley & Sons, New York, 1979.
5. S. G. Tzafestas, *Walsh Functions in Signal and Systems Analysis and Design*, Van Nostrand Company, 1985.
6. F. M. Gardner, Acquisition of phaselock, *Conference Record of the International Conference on Communications*, ICC'76 vol. I, pp. 10.1–10.5, June 1976.
7. D. Richman, Color carrier reference phase synchronization accuracy in NTSC color television, *Proc. IRE*, 42, 106–133, 1954.
8. F. M. Gardner, Properties of frequency difference detector, *IEEE Trans. Commun.*, COM-33, 131–138, 1985.
9. D. G. Messerschmitt, Frequency detectors for PLL acquisition in timing and carrier recovery, *IEEE Trans. Commun.*, COM-27, 1288–1295, 1979.
10. R. B. Lee, Timing recovery architecture for high speed data communication system, Master thesis, Department of Electrical Engineering, National Taiwan University, Taiwan, ROC, 1993.
11. M.-T. Shiue, Transceiver VLSI design for high speed local access modems, PhD thesis, Department of Electrical Engineering, National Central University, Chung-Li, Taiwan, ROC, 1998.
12. M. Bazes, A novel precision MOS synchronous delay lines, *IEEE J. Solid-State Circuits*, 20, 1265–1271, 1985.
13. M. G. Johnson and E. L. Hudson, A variable delay line PLL for CPU-coprocessor synchronization, *IEEE J. Solid-State Circuits*, 23, 1218–1223, 1988.
14. B. Kim, T. C. Weigandt, and P. R. Gray, PLL/DLL systems noise analysis for low jitter clock synthesizer design, *IEEE International Symposium on Circuits and Systems*, ISCAS'94, 4, pp. 31–34, 1994.
15. M. V. Paemel, Analysis of a charge-pump PLL: A new model, *IEEE Trans. Commun.*, COM-42, 131–138, 1994.

16. F. M. Gardner, Charge-pump phase-locked loops, *IEEE Trans. Commun*., COM-28, 1849–1858, 1980.
17. F. M. Gardner, Phase accuracy of charge pump PLL's, *IEEE Trans. Commun*., COM-30, 2362–2363, 1982.
18. T. C. Weigandt, B. Kim, and P. R. Gray, Analysis of timing recovery jitter in CMOS ring oscillator, *IEEE International Symposium on Circuits and Systems*, ISCAS'94, 4, pp. 27–30, 1994.
19. T. H. Lee and J. F. Bulzacchelli, A 155-MHz clock recovery delay- and phase-locked loop, *IEEE J. Solid-State Circuits*, 27, 1736–1746, 1992.
20. M. P. Flyun and S. U. Lidholm, A 1.2 μm CMOS current-controlled oscillator, *IEEE J. Solid-State Circuits*, 27, 982–987, 1992.
21. S. K. Enam and A. A. Abidi, NMOS IC's for clock and data regeneration in gigabit-per-second optical-fiber receivers, *IEEE J. Solid-State Circuits*, 27, 1763–1774, 1992.
22. M. Horowitz et al., PLL design for a 500 MB/s Interface, *IEEE International Solid-State Circuit Conference/Digest of Technical Papers*, pp. 160–161, 1993.
23. A. Pottbacker and U. Langmann, An 8 GHz silicon bipolar clock-recovery and data-regenerator IC, *IEEE J. Solid-State Circuits*, 29, 1572–1751, 1994.
24. B. P. Lathi, *Modern Digital and Analog Communication System*, HRW, Philadelphia, 1989.
25. J. A. Bellisio, A new phase-locked loop timing recovery method for digital regenerators, *IEEE Int. Comm. Conf. Rec*., 1, 10.17–10.20, 1976.
26. B. Razavi, A 2.5-Gb/s 15-mW clock recovery circuit, *IEEE J. Solid-State Circuits*, 31, 472–480, 1996.
27. R. J. Baumert, P. C. Metz, M. E. Pedersen, R. L. Pritchett, and J.A. Young, A monolithic 50–200 MHz CMOS clock recovery and retiming circuit, *Proceedings of the IEEE Custom Integrated Circuits Conference*, 14.5.5–14.5.4, San Diego, CA, 1989.
28. B. Lai and R. C. Walker, A monolithic 622 Mb/s clock extraction data retiming circuit, *IEEE International Solid-State Circuit Conference/Digest of Technical Papers*, pp. 144–145, 1991.
29. B. Kim, D. M. Helman, and P. R. Gray, A 30 MHz hybrid analog/digital clock recovery circuit in 2-μm CMOS, *IEEE J. Solid-State Circuits*, 25, 1385–1394, 1990.
30. R. Jaffe and E. Rechtin, Design and performance of phase-lock circuits capable of near-optimal performance over a wide range of input signal and noise levels, *IRE Trans. Inf. Theory*, IT-1, 66–76, 1955.
31. B. Chun, Y. H. Lee, and B. Kim, Design of variable loop gains of dual-loop DPLL, *IEEE Trans. Commun*., COM-45, 1520–1522, 1997.
32. J. Lee and B. Kim, A low-noise fast-lock phase-locked loop with adaptive bandwidth control, *IEEE J. Solid-State Circuits*, 35, 1137–1145, 2000.
33. J. Dunning, et al., An all-digital phase-locked loop with 50-cycle lock time suitable for high-performance microprocessors, *IEEE J. Solid-State Circuits*, 30, 412–422, 1995.
34. J. G. Maneatis, Low-jitter process-independent DLL and PLL based on self-biased techniques, *IEEE J. Solid-State Circuits*, 31, 1723–1732, 1996.
35. S. Sidiropoulos, et al., Adaptive bandwidth DLLs and PLLs using regulated supply CMOS buffers, in *IEEE Symp. VLSI Circuits Dig. Tech. Papers*, June 2000, pp. 124–127.
36. J. Kim and M. A. Horowitz, Adaptive-supply serial links with sub-1 V operation and per-pin clock recovery, *IEEE J. Solid-State Circuits*, 37, 1403–1413, 2002.
37. J. Kim, et al., Design of CMOS adaptive-bandwidth PLL/DLLs: A general approach, *IEEE Trans. Circuits Systems-II*, 50, 860–869, 2003.
38. http://www.rambus.com/products/innovationslicensing/innovations/dllpll.aspx

6

Synthesis of Reactance Pulse-Forming Networks

6.1 Introduction **6**-1
6.2 Networks Forming Quasi-Rectangular Output Pulses **6**-2
Quasi-Rectangular Output Pulse and Its Laplace Transform • Realization Requirements • Second Approximation Step: Approximation for Realization • Pulse-Forming Networks with Non-Delayed Output Pulses • Pulse-Forming Networks with Delayed Output Pulse
6.3 Transfer Functions of Wideband Amplifiers **6**-15
Parameters of the Step Response and Its Laplace Transform • Transfer Function Approximation • Example of Transfer Function Design • Tabulated Results
6.4 Forming a Sinusoidal Pulse **6**-19
Required Transfer Function • Approximation for Realization • Example
6.5 Summary **6**-24
References **6**-25

Igor M. Filanovsky
University of Alberta

6.1 Introduction

Linear lumped reactance parameter networks loaded by resistors can be used to form pulses of different shapes [1,2]. The most known applications [2] require the networks that form pulses of the so-called quasi-rectangular shape. The network excitation is usually a step voltage, and it is required that the network output response (in this case it is the step response) is a pulse of finite duration. In a realizable network, the fronts of this pulse should have ''rounded'' corners, and the slopes of the fronts should also be finite. Below this is done using a semi-period of $\sin^2 t$ function.

These pulse-forming networks will be considered first, and it will be shown that the required result may be obtained using different input excitations and using, of course, different networks. The pulse of the same shape can be obtained when the excitation is a step, an impulse or even a sinusoidal function. The last result looks surprising for linear networks, yet the reader should understand that the output pulse is shaped during the transient period, when the sinusoidal voltage is turning on. When the transient is over, the steady-state response of the network is zero.

This reconsideration of the old problem establishes a certain difference between synthesis in time domain and frequency domain. The output signal time-domain approximation, in case of pulse-forming

networks, should also result in a realizable transfer function. But this synthesis results in many solutions not only because of many possible realizations of the same transfer function. In addition, one may change input excitation (even though in practice the choice of input excitations is limited), then find modified transfer function, and then synthesize new networks. This modification of input signal is not used in the frequency-domain synthesis where it is assumed that the input excitation is the sinusoidal signal.

The method of approximation proposed here gives new and useful results even for the case of the step voltage input excitation and quasi-rectangular shape of the output response. It is shown that if the output pulse is delayed with respect to the input step, then it is possible to shape the output pulse so that its amplitude will be higher than the amplitude of the applied input step voltage. The increase of amplitude is usually connected with the idea of using a transformer, yet the pulse-shaping circuit may provide transformerless change of amplitude.

The approximation developed here for pulse forming may also be applied to the synthesis of wideband amplifier transfer functions. Unfortunately, an initial attempt of using $\sin^2 t$ for this purpose [3] did not bring any general results, and the required transfer functions were found numerically [4]. But even this attempt was forgotten, and the transfer functions of the wideband amplifiers were found indirectly. They were obtained starting from the frequency domain and investigating the time-domain response of the filter transfer functions. It was found that the step responses of the Bessel filters do not have overshoot. It happened that exactly this transient response is required in realization of the wideband amplifiers. Yet, as shown here, using $\sin^2 t$ in the approximation of the impulse response allows one to find the realizable transfer functions directly, on the basis of requirements formulated in the time domain to the step response.

Finally, the considered approximation method allows one to find realizable transfer functions for the networks with the step or impulse response representing the sinusoidal pulse of finite duration, and with the pulse envelope that can be of an arbitrary shape. As an example, we consider the case when this envelope is also of sinusoidal shape. The pulses of these shapes find applications in ultra-wideband transmission networks and wavelet signal analysis. These pulse-shaping networks are usually realized by active networks. Yet, the fact that at least one physically realizable network is available is important for using of algorithms, which are able to improve an approximation and find a transfer function that is better suitable for the chosen realization method.

Many of the results given here can be found in [5]. Considering that this source of information is not easily accessible, a sincere effort is done to summarize all available material.

6.2 Networks Forming Quasi-Rectangular Output Pulses

In this section, we consider all procedures required in the synthesis of pulse-forming networks [5,6]. They start by approximation of the output pulse in the time domain and finding the Laplace transform of this pulse. Then, we review the requirements imposed on the realizable transfer function. After that we consider the second round of approximation, when the Laplace transform is approximated by an algebraic ratio obtained from the spectral representation of the Laplace transform. Finally, we give examples of realization. Some of these procedures will be omitted in parts dedicated to the wideband amplifier transfer functions and forming of the sinusoidal pulses.

6.2.1 Quasi-Rectangular Output Pulse and Its Laplace Transform

Assume that it is required to find reactance networks forming a quasi-rectangular pulse for three cases of input excitation, $v_i(t)$, namely, the unit impulse, the unit step, and the sinusoid with unit amplitude. The ratio of two polynomials approximating the Laplace transform of the output response will be the same for all three cases. In case of impulse excitation, $v_i(t) = \delta(t)$, this ratio is directly equal to the transfer function of the pulse-forming network. In case of $v_i(t) = u(t)$, a unit step voltage, the required transfer function is obtained by multiplication of the approximating ratio by s, and in case of $v_i(t) = \sin \Omega t$ this ratio is multiplied by $(s^2 + \Omega^2)$.

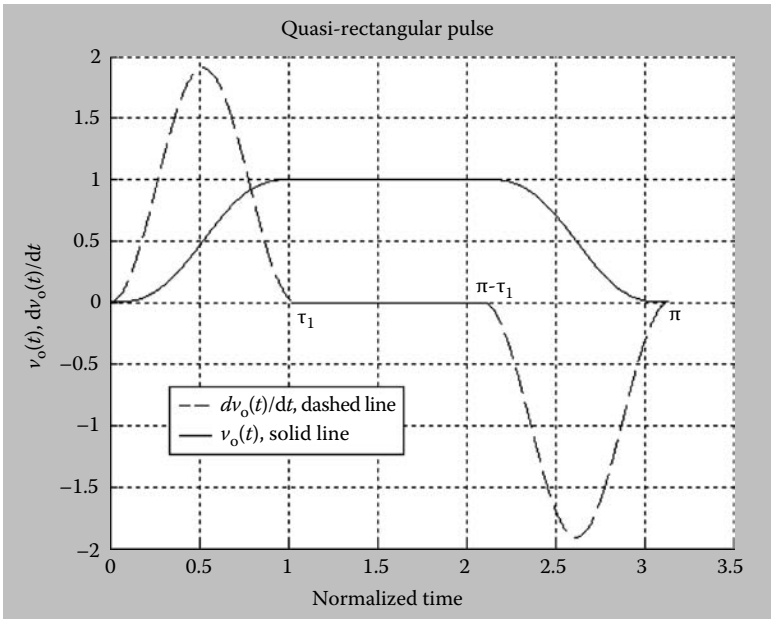


FIGURE 6.1 Quasi-rectangular pulse.

The required output response, $v_o(t)$, is a symmetric pulse (Figure 6.1) of the normalized duration $\tau = \pi$ for all considered cases of excitation. Assume that the derivative, dv_o/dt, of this response is approximated by one semi-period of $\sin^2(\pi t/\tau_1)$ function for the initial front of the pulse, and one delayed negative semi-period of $\sin^2(\pi t/\tau_1)$ for the rear front. The duration of each front is equal to τ_1. Then dv_o/dt is described by the following system of equations:

$$\frac{dv_o}{dt} = \begin{cases} A\sin^2(\pi t/\tau_1) & 0 < t < \tau_1 \\ 0 & \tau_1 < t < (\tau - \tau_1) \\ -A\sin^2[\pi(t - \tau + \tau_1)/\tau_1] & (\tau - \tau_1) < t < \tau \\ 0 & t > \tau = \pi \end{cases}. \tag{6.1}$$

The amplitude, A, will be obtained from the normalization condition

$$A\int_0^{\tau_1} \sin^2(\pi t/\tau_1)dt = 1. \tag{6.2}$$

This gives $A = 2/\tau_1$. This value of A is also equal to the maximal slew rate of the fronts, so that the approximate rise and fall times can be evaluated as

$$\tau_r = \tau_f = 1/A = \tau_1/2. \tag{6.3}$$

These rise and fall times are close to ones defined by 0.1 and 0.9 levels of the output pulse.

The Laplace transform of $(2/\tau_1)\sin^2(\pi t/\tau_1)$ function is equal to [7]

$$F_0(s) = \frac{1}{\tau_1}\frac{4(\pi/\tau_1)^2}{s\left[s^2 + 4(\pi/\tau_1)^2\right]}. \tag{6.4}$$

Using the shift theorem [8], one finds that the Laplace transform of the sine-squared pulse will be

$$F_1(s) = \frac{1}{\tau_1} \frac{4(\pi/\tau_1)^2}{s\left[s^2 + 4(\pi/\tau_1)^2\right]} (1 - e^{-s\tau_1}). \tag{6.5}$$

The derivative dv_0/dt includes two (one positive, another negative) sine-squared pulses shifted with respect to each other by $\tau - \tau_1$. Using the shift theorem again, and considering that

$$\mathbf{L}[dv_0/dt] = sV_o(s), \tag{6.6}$$

where $V_o(s)$ is the Laplace transform of $v_o(t)$, one obtains that

$$V_o(s) = \frac{1}{\tau_1} \frac{4(\pi/\tau_1)^2}{s^2\left[s^2 + 4(\pi/\tau_1)^2\right]} (1 - e^{-s\tau_1}) \left[1 - e^{-s(\pi - \tau_1)}\right]. \tag{6.7}$$

This $V_o(s)$ may be considered as an impulse response of a nonrealizable (using lumped parameters) transfer function. But if, following [5], its real, $\mathrm{Re}V_o(j\omega)$, and imaginary, $\mathrm{Im}V_o(j\omega)$, parts are expanded in infinite products, then the finite number of terms in these products may be used for approximation of $\mathrm{Re}V_o(j\omega)$ and $\mathrm{Im}V_o(j\omega)$. This further approximation allows one to obtain a realizable transfer function $H_0(s)$ of the pulse-forming network excited by the input unit impulse. As one will see below, the obtained transfer function includes zeros on the $j\omega$-axis only, and the denominator order may be controlled.

Then, if $v_i(t)$ is a step function, the required transfer function is equal to $H_1(s) = sH_0(s)$. If $v_i(t)$ is sin Ωt, the required transfer function is $H_2(s) = (s^2 + \Omega^2)H_0(s)$.

6.2.2 Realization Requirements

We remind here the requirements that should be imposed on the algebraic ratio if it represents the transfer function of a reactance network loaded by resistor [9,10].

Let $F(s)$ be the Laplace transform of network impulse response, and $F(j\omega) = P(\omega) + jQ(\omega)$. One can find a realizable network if $F(s)$ allows the approximation

$$H_a(s) = \frac{N(s)}{D(s)} = \frac{A_1 s^\mu N_1(s)}{D_2(s) + A_2 s D_1(s)}. \tag{6.8}$$

Here $H_a(s)$ is the ratio of two polynomials, and deg $N(s) \le$ deg $D(s)$, $N_1(s)$, $D_1(s)$, and $D_2(s)$ are even polynomials with zeros on the $j\omega$-axis only, that is, $N_1(s) = \prod_{l=1}^{\nu} \left(s^2 + c_l^2\right)$, $D_2(s) = \prod_{k=1}^{n} \left(s^2 + a_k^2\right)$, $sD_1(s) = s\prod_{k=1}^{n-1} \left(s^2 + b_k^2\right)$, and a_k and b_k are alternating (for odd denominator polynomials both product superscripts in $D_1(s)$ and $D_2(s)$ should be equal). A_1 and A_2 are the positive real constants. Indeed, the function (Equation 6.8) allows then realization as a transfer function of a reactance network loaded by resistor [9,10].

The conditions that should be imposed on real and imaginary parts of realizable $H_a(s)$ are obtained the following way. To avoid separate discussions for even and odd μ let us consider the function

$$H_{a1}(s) = s^{-\mu} H_a(s) = \frac{A_1 N_1(s)}{D_2(s) + A_2 s D_2(s)}. \tag{6.9}$$

Then one can write that

$$\mathrm{Re}\, H_{a1}(j\omega) = \frac{A_1 \prod_{l=1}^{\nu} \left(c_l^2 - \omega^2\right) \prod_{k=1}^{n} \left(a_k^2 - \omega^2\right)}{D_2^2(-\omega^2) + A_2^2 \omega^2 D_1(-\omega^2)} \tag{6.10}$$

and

$$\operatorname{Im} H_{a1}(j\omega) = -\frac{A_1 A_2 \omega \prod_{l=1}^{\nu}\left(c_l^2 - \omega^2\right)\prod_{k=1}^{n-1}\left(b_k^2 - \omega^2\right)}{D_2^2(-\omega^2) + A_2^2\omega^2 D_1(-\omega^2)}. \tag{6.11}$$

Hence, if $F(s) = s^{\mu_1}F_1(s)$ and $F_1(j\omega) = P_1(\omega) + jQ_1(\omega)$, then (1) both $\mathrm{Re}F_1(j\omega)$ and $\mathrm{Im}F_1(j\omega)$ should be approximated by the ratios so that each of them includes 2ν simple zeros located at $\omega = \pm\gamma_k$, (2) the approximating ratio for $\mathrm{Re}F_1(j\omega)$ includes, in addition, $2n$ simple zeros located at $\omega = \pm\alpha_k$, (3) the approximating ratio for $\mathrm{Im}F_1(j\omega)$ includes, in addition, $2(n-1)$ simple zeros located at $\omega = \pm\beta_k$, and (4) α_k and β_k should alternate. Then the approximating ratio

$$H_a(s) = \frac{A_1 s^{\mu_1}\prod_{l=1}^{\nu}\left(s^2 + \gamma_l^2\right)}{\prod_{k=1}^{n}\left(s^2 + \alpha_k^2\right) + A_2 s\prod_{k=1}^{n-1}\left(s^2 + \beta_k^2\right)}, \tag{6.12}$$

where $\mu_1 = \mu$, and $a_k = \alpha_k$, $b_k = \beta_k$, and $c_l = \gamma_l$, may be used as a transfer function for the network realization. The constants A_1 and A_2 can be calculated equating $F(s)$ and $H_a(s)$ at two points of the s-plane (the most convenient points are usually $s = 0$ and $s = j$, the final choice may be decided by computer calculations).

One can see that if $\deg N(s) \le \deg D(s) - 2$, the multiplication of the numerator in Equation 6.12 by s or by $(s^2 + \Omega^2)$ does not violate the realization condition.

Representing thus obtained $H_a(s)$, for example, as

$$H_a(s) = \frac{[A_1 s^{\mu} N_1(s)]/D_2(s)}{\{[A_2 s D_1(s)]/D_2(s)\} + 1} = -\frac{y_{21}(s)}{y_{22}(s) + 1}, \tag{6.13}$$

one finds $y_{21}(s) = -A_1 s^{\mu} N_1(s)/D_2(s)$ and $y_{22}(s) = A_2 s N_1(s)/D_2(s)$. These two parameters are sufficient to realize the network [9,10] within a constant multiplier for $y_{21}(s)$.

6.2.3 Second Approximation Step: Approximation for Realization

Now we find the approximation to the output response Laplace transform (Equation 6.7). This step is leading to realizable transfer functions. We rewrite $V_o(s)$ as

$$V_o(s) = \frac{16(\pi/\tau_1)^2}{\tau_1 s^2\left[s^2 + 4(\pi/\tau_1)^2\right]} \cdot e^{-\frac{s\pi}{2}} \cdot \sinh\left(\frac{s\tau_1}{2}\right) \cdot \sinh\left[\frac{s(\pi - \tau_1)}{2}\right]. \tag{6.14}$$

Substituting $s = j\omega$ one can find that

$$\mathrm{Re}V_o(j\omega) = \frac{16(\pi/\tau_1)^2}{\tau_1\omega^2\left[4(\pi/\tau_1)^2 - \omega^2\right]} \cdot \cos\left(\frac{\omega\pi}{2}\right) \cdot \sin\left(\frac{\omega\tau_1}{2}\right) \cdot \sin\left[\frac{\omega(\pi - \tau_1)}{2}\right] \tag{6.15}$$

and

$$\mathrm{Im}V_a(j\omega) = -\frac{16(\pi/\tau_1)^2}{\tau_1\omega^2\left[4(\pi/\tau_1)^2 - \omega^2\right]} \cdot \sin\left(\frac{\omega\pi}{2}\right) \cdot \sin\left(\frac{\omega\tau_1}{2}\right) \cdot \sin\left[\frac{\omega(\pi - \tau_1)}{2}\right]. \tag{6.16}$$

Taking into consideration that [11]

$$\sin x = x\prod_{k=1}^{\infty}\left(1-\frac{x^2}{k^2\pi^2}\right) \tag{6.17}$$

and

$$\cos x = \prod_{k=0}^{\infty}\left(1-\frac{x^2}{(2k+1)^2\pi^2}\right), \tag{6.18}$$

one can see that $\mathrm{Re}V_o(j\omega)$ has real zeros at $\alpha_k = \pm(2k+1)$, where $k = 0, 1, 2, \ldots$, and $\mathrm{Im}V_a(j\omega)$ has zeros at $\beta_k = \pm 2k$, where $k = 0, 1, 2, \ldots$ as well. These zeros are alternating on ω-axis. In addition, both $\mathrm{Re}V_o(j\omega)$ and $\mathrm{Im}V_o(j\omega)$ have common zeros located at $\gamma_k = \pm 2\pi\ k$, where $k = 2, 3, 4, \ldots$ (the zeros for $k = 0, 1$ are canceling the poles), and $\gamma_k = \pm 2\pi k/(\pi - \tau_1)$, where $k = 1, 2, 3 \ldots$ (the zero for $k = 0$ is canceling the pole). Then, in accordance with the previous part, to approximate $V_a(s)$ one may use the algebraic ratio

$$H_a(s) = \frac{N(s)}{D(s)} = \frac{A_1\prod_{k=2}^{\nu_1}\left[s^2+(2\pi k/\tau_1)^2\right]\prod_{k=1}^{\nu_2}\left[s^2+(2\pi k/(\pi-\tau_1))^2\right]}{\prod_{k=0}^{n_1}\left[s^2+((2k+1)^2\right]+A_2 s\prod_{k=1}^{n}(s^2+4k^2)} \tag{6.19}$$

To have the approximation with sufficient number of terms in the products, it is rational to find the spectral function

$$|V_o(j\omega)| = \left|\frac{16(\pi/\tau_1)^2\sin(\omega\tau_1/2)\sin[\omega(\pi-\tau_1)/2]}{\tau_1\omega^2\left[4(\pi/\tau_1)^2-\omega^2\right]}\right| \tag{6.20}$$

and evaluate the spectrum bandwidth ω_m. The zeros included in the approximating products should be located in the spectrum bandwidth, and they will determine the subscripts in the products of Equation 6.21. The following two possibilities should be considered for n_1 and n. When $n_1 = n - 1$, then the denominator degree is odd, when $n_1 = n$ then this degree is even. If ω_m is chosen, then, in case of the odd degree $2n \approx \omega_m$ (i.e., $n \approx \omega_m/2$), and in case of the even degree $2n_1 + 1 \approx \omega_m$ (i.e., $n_1 \approx \omega_m - 0.5$). Hence, the odd degree allows one to obtain the approximating transfer function with more "dense" location of real and imaginary part zeros in the spectrum bandwidth, which usually results in a better approximation in the time domain as well.

From the other side, to obtain the simplest transfer function one has to choose $\nu_1 = 1$. Then, the first product in the numerator is equal to unit and will not include any multipliers. Considering that $(2\pi/\tau_1)$ represents now the upper approximation frequency, one can find the other superscripts in the approximating products from the approximate equalities

$$\begin{cases}\dfrac{2\pi}{\tau_1}\approx\dfrac{2\pi\nu_2}{\pi-\tau_1}\\[2ex]\dfrac{2\pi}{\tau_1}\approx 2n \text{ or } (2n_1+1)\end{cases} \tag{6.21}$$

that give $\nu_2 \approx (\pi - \tau_1)/\tau_1$ and $n \approx \dfrac{\pi}{\tau_1}$; $n_1 \approx \dfrac{\pi}{\tau_1} - \dfrac{1}{2}$. This results in

$$H_a(s) = \frac{A_1\prod_{k=1}^{\nu_2}\left[s^2+\left(\frac{2\pi k}{\pi-\tau_1}\right)^2\right]}{\prod_{k=0}^{n_1}\left[s^2+(2k+1)^2\right]+A_2 s\prod_{k=1}^{n}(s^2+4k^2)}. \tag{6.22}$$

The constants A_1 and A_2 can be found equating $V_o(s)$ given by Equation 6.7 and this $H_a(s)$ at two points of the s-plane. Using Equation 6.13 one can then find the realizable $y_{22}(s)$ and $y_{21}(s)$ required for the network synthesis.

6.2.4 Pulse-Forming Networks with Non-Delayed Output Pulses

As examples, we synthesize the pulse-forming networks that should have the output response, which is the quasi-rectangular pulse with duration of fronts equal to $\tau_1 = \pi/3$, so that the rise and fall times are $\pi/6$ each. To obtain the simplest network, one has to put $\nu_2 = 2$, $n = 3$, and $n_1 = 2$. Then $V_o(s)$ for the case of $v_i = \delta(t)$ will be approximated by the ratio

$$H_a(s) = \frac{A_1(s^2+9)(s^2+36)}{[(s^2+1)(s^2+9)(s^2+25) + A_2 s(s^2+4)(s^2+16)(s^2+36)]}. \tag{6.23}$$

Equating Equations 6.7 and 6.23 at $s = 0$ and $s = j$ one can find that $A_1 = 1.453$ and $A_2 = 0.152$. Hence, the network transfer function for this case is equal to

$$H_0(s) = \frac{1.453(s^2+9)(s^2+36)}{[(s^2+1)(s^2+9)(s^2+25) + 0.152s(s^2+4)(s^2+16)(s^2+36)]}. \tag{6.24}$$

The impulse response, $h_o(t)$, of this transfer function is shown in Figure 6.2 (the initial output pulse, $v_o(t)$, is also shown for comparison).

If $v_i = u(t)$ then the network transfer function for this case of excitation is equal to

$$H_1(s) = \frac{1.453s(s^2+9)(s^2+36)}{[(s^2+1)(s^2+9)(s^2+25) + 0.152s(s^2+4)(s^2+16)(s^2+36)]}. \tag{6.25}$$

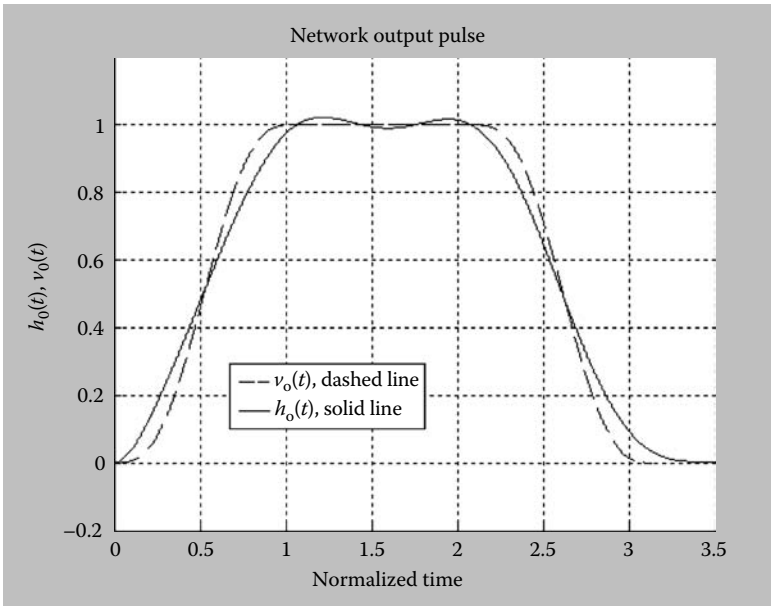


FIGURE 6.2 Pulse-forming network output response.

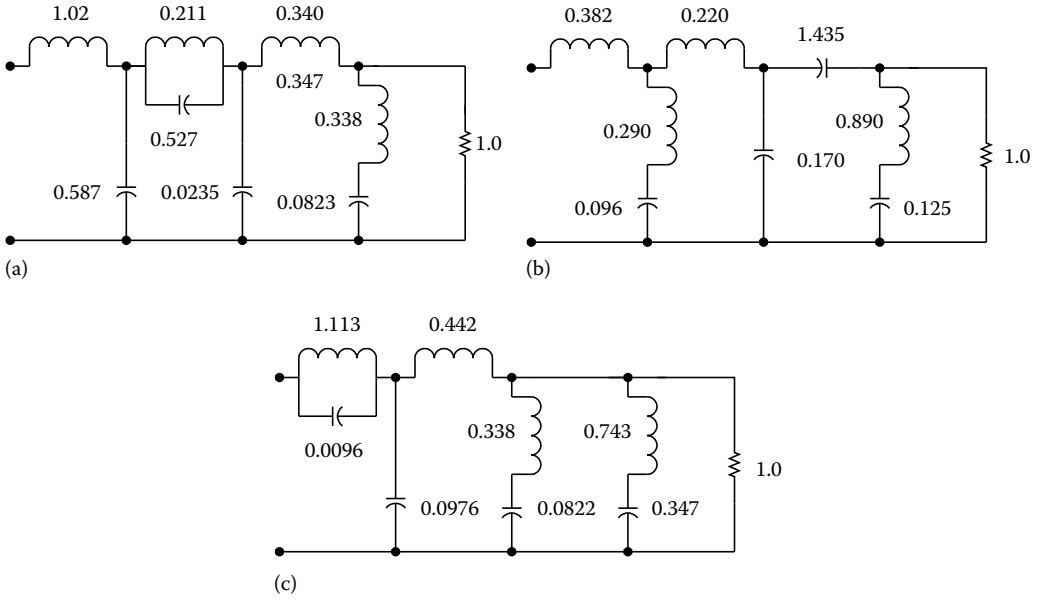


FIGURE 6.3 Pulse-forming networks: (a) $v_i = \delta(t)$, (b) $v_i = u(t)$, and (c) $v_i = \sin \Omega t$.

Finally, assume that $v_i = \sin 4t$. Then the zeros that are due to the input signal will coincide with one pair of zeros in the denominator odd part. This will simplify the realization. The transfer function for this case of excitation is equal to

$$H_2(s) = \frac{1.453(s^2+9)(s^2+16)(s^2+36)}{[(s^2+1)(s^2+9)(s^2+25)+0.152s(s^2+4)(s^2+16)(s^2+36)]}. \tag{6.26}$$

The realizations of the functions (Equations 6.24 through 6.26) were obtained dividing their numerators and denominators by $(s^2+1)(s^2+9)(s^2+25)$. This gives

$$y_{22}(s) = \frac{0.152s(s^2+4)(s^2+16)(s^2+36)}{(s^2+1)(s^2+9)(s^2+25)}, \tag{6.27}$$

common for all three networks. Then this parameter was realized by standard procedures [9,10] taking into consideration the location of $y_{12}(s)$ zeros for each network. The results of these realizations are shown in Figure 6.3.

One can also verify that for the network of Figure 6.3b, the voltage source efficiency coefficient representing the ratio of the output pulse amplitude to input step amplitude is equal to 0.743, that is, nearly 1.5 times higher than with the usual approach [1,2].

If the required duration of pulse is τ_0 s, and the load resistor is R ohms, then each value of inductance should be multiplied by $\tau_0 R/\pi$ and the value of each capacitor by $\tau_0/(R\pi)$.

6.2.5 Pulse-Forming Networks with Delayed Output Pulse

The above-mentioned voltage source efficiency coefficient increases when the output pulse is delayed [12]. Yet, this delay should be used judiciously. The increase of source efficiency results in a more complicated circuit. In addition, the rise and fall times increase, that is, slew rate of the fronts

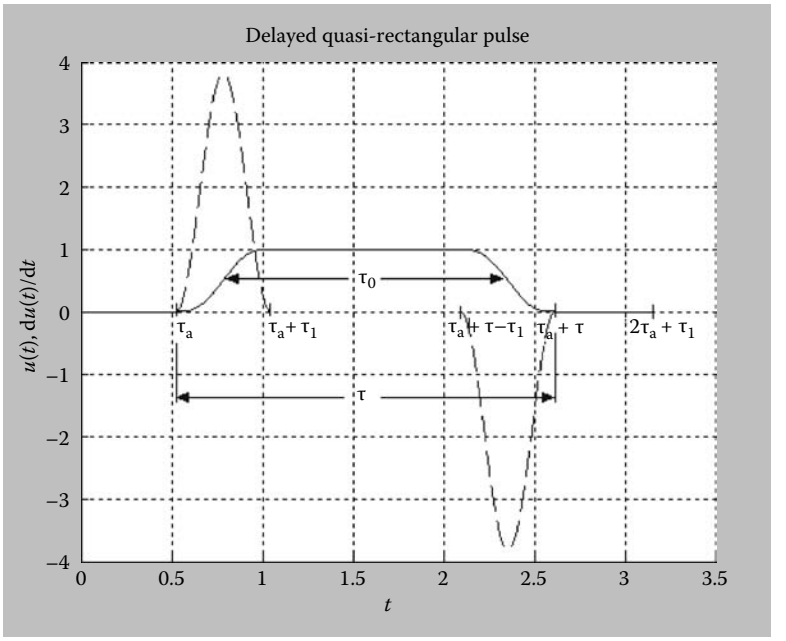


FIGURE 6.4 Delayed quasi-rectangular pulse.

deteriorates. It is possible to prepare a table that helps to visualize a possible trade-off between the delay, the circuit complexity, and the voltage source efficiency coefficient. The calculations given below follow the above-developed pattern.

Let the required output response be a delayed symmetric pulse $u(t)$ (Figure 6.4). Assume that the derivative, $\mathrm{d}u/\mathrm{d}t$, of this response is described by delayed positive and delayed negative semi-periods of sine-squared function. These semi-periods determine the duration of fronts that is equal to τ_1. Then $\mathrm{d}u/\mathrm{d}t$ can be described by the following equations:

$$\frac{\mathrm{d}u}{\mathrm{d}t} = \begin{cases} 0 & 0 \le t < \tau_a \\ A\sin^2\left[\frac{\pi}{\tau_1}(t-\tau_a)\right] & \tau_a \le t < \tau_a + \tau_1 \\ 0 & \tau_a + \tau_1 \le t < \tau_a + \tau - \tau_1 \\ -A\sin^2\left[\frac{\pi}{\tau_1}(t+\tau_1-\tau-\tau_a)\right] & \tau_a + \tau - \tau_1 \le t < \tau_a + \tau \\ 0 & t > \tau_a + \tau \end{cases} \tag{6.28}$$

The normalized time is defined by the condition

$$2\tau_a + \tau = \pi. \tag{6.29}$$

To obtain the normalized magnitude of unity for the output pulse amplitude, one has to choose the value of A from the equation

$$A\int_{\tau_a}^{\tau_a+\tau_1} \sin^2\left[\frac{\pi}{\tau_1}(t-\tau_a)\right]\mathrm{d}t = 1. \tag{6.30}$$

This gives

$$A = 2/\tau_1. \tag{6.31}$$

This value is again equal to the maximal slew rate of the fronts, and the approximate rise and fall times can be evaluated as

$$\tau_r = \tau_f = 1/A = \tau_1/2. \tag{6.32}$$

The value of $\tau_0 = \tau - \tau_1$ will be considered as the pulse duration at the level of 0.5.

We repeat here, for convenience, the Laplace transform of $(2/\tau_1)\sin^2(\pi t/\tau_1)$ [7] that is equal to

$$F_0(s) = \frac{1}{\tau_1}\frac{4(\pi/\tau_1)^2}{s\left[s^2 + 4(\pi/\tau_1)^2\right]}. \tag{6.33}$$

The derivative du/dt includes two (one positive, another negative) delayed sine-squared pulses shifted with respect to each other by $\tau_0 = \tau - \tau_1$. Using the shift theorem [8], and, considering that $\mathbf{L}[du/dt] = sU(s)$, where $U(s)$ is the Laplace transform of $u(t)$, one obtains that

$$U(s) = \frac{4(\pi/\tau_1)^2}{\tau_1}\frac{e^{-s\tau_a}(1 - e^{-s\tau_1})\left[1 - e^{-s(\tau-\tau_1)}\right]}{s^2\left[s^2 + 4(\pi/\tau_1)^2\right]}. \tag{6.34}$$

This result can be rewritten as

$$U(s) = \frac{16(\pi/\tau_1)^2}{\tau_1}\frac{e^{-s\left(\tau_a+\frac{\tau}{2}\right)}\sinh\left(\frac{s\tau_1}{2}\right)\sinh\left[\frac{s(\tau-\tau_1)}{2}\right]}{s^2\left[s^2 + 4(\pi/\tau_1)^2\right]}. \tag{6.35}$$

For the normalized time defined by Equation 6.29, one can rewrite this result as

$$U(s) = \frac{16(\pi/\tau_1)^2}{\tau_1}\frac{e^{-s\frac{\pi}{2}}\sinh\left(\frac{s\tau_1}{2}\right)\sinh\left[\frac{\xi\pi s}{2}\right]}{s^2\left[s^2 + 4(\pi/\tau_1)^2\right]}, \tag{6.36}$$

where $\xi = \frac{1}{1+(2\tau_a+\tau)/\tau_0} < 1$. Using the expansion [11] of

$$\sinh(x) = x\prod_{k=1}^{\infty}\left(1 + \frac{x^2}{4k^2\pi^2}\right) \tag{6.37}$$

and approximating

$$\sinh\left(s\frac{\tau_1}{2}\right) \approx \frac{s\tau_1}{2}\left(1 + \frac{s^2\tau_1^2}{4\pi^2}\right) \tag{6.38}$$

and

$$\sinh\left(\frac{\xi\pi s}{2}\right) \approx \frac{\xi\pi s}{2}\prod_{l=1}^{\nu}\left[1 + \frac{s^2}{(2l/\xi)^2}\right] \tag{6.39}$$

(the choice of ν is discussed below), one can write

$$U(s) \approx Ce^{-s\frac{\pi}{2}} \prod_{l=1}^{\nu} \left[s^2 + \left(\frac{2l}{\xi} \right)^2 \right], \tag{6.40}$$

where $C = (\xi\pi)^{2\nu+1} / \left(4^{\nu} \prod_{l=1}^{\nu} l^2 \right)$. The approximation (Equation 6.38) defines the maximum frequency, ω_m, where Equation 6.40 is valid on the $j\omega$-axis, namely,

$$(2\pi)/\tau_1 \leq \omega_m < (4\pi)/\tau_1 \tag{6.41}$$

(the calculations show that using $(3\pi)/\tau_1$ for the right-hand side is usually sufficient).

In this problem, the output pulse $u(t)$ is usually the result of application of a step voltage to the network input. Then, if the Laplace transform of this input voltage is written as

$$U_1(s) = 1/(Ks), \tag{6.42}$$

then the realized value of K (it will be defined below) gives us the voltage source efficiency coefficient.

The ratio $U(s)/U_1(s)$ is a nonrealizable transfer function. Yet, using the previously developed approach, it is possible to find further approximation to Equation 6.40 so that the resulting ratio will be realizable.

We calculate $\mathrm{Re}U(j\omega)$ and $\mathrm{Im}U(j\omega)$ for the approximation (Equation 6.40). Substituting $s = j\omega$ one finds that

$$\mathrm{Re}U(j\omega) \approx C \cos\left(\omega \frac{\pi}{2} \right) \prod_{l=1}^{\nu} \left[\left(\frac{2l}{\xi} \right)^2 - \omega^2 \right] \tag{6.43}$$

and

$$\mathrm{Im}U(j\omega) \approx -C \sin\left(\omega \frac{\pi}{2} \right) \prod_{l=1}^{\nu} \left[\left(\frac{2l}{\xi} \right)^2 - \omega^2 \right]. \tag{6.44}$$

Taking into consideration Equations 6.18 and 6.19, one can see that $\mathrm{Re}U(j\omega)$ has real zeros at $\alpha_k = \pm(2k+1)$ with $k = 0, 1, 2, \ldots$ and $\mathrm{Im}U(j\omega)$ has zeros at $\beta_k = \pm 2k$ with $k = 0, 1, 2, \ldots$ as well. These zeros are alternating (or interlacing). In addition, both $\mathrm{Re}U(j\omega)$ and $\mathrm{Im}U(j\omega)$ have common zeros located at $\gamma_l = \pm 2l/\xi$ with $l = 1, 2, \ldots, \nu$. Then, in accordance with the previous part, to approximate $U(s)$ in the bandwidth defined by Equation 6.41, one can use the algebraic ratio

$$U_a(s) = \frac{A_1 \prod_{l=1}^{\nu} [s^2 + (2l/\xi)^2}{\prod_{k=0}^{n_1} \left[s^2 + (2k+1)^2 \right] + A_2 s \prod_{k=1}^{n} (s^2 + 4k^2)}. \tag{6.45}$$

The coefficients A_1 and A_2 may be found equating $U_a(s)$ given by Equation 6.45 to $U(s)$ given by Equation 6.36 at two points of the s-plane.

The Equation 6.41 defines the maximum approximation bandwidth ω_m for Equation 6.45 as well. To obtain simpler networks, it is better to be closer to the left-hand side of this inequality. Then, one can find ν from the approximate condition

$$2\nu/\xi \approx 2\pi/\tau_1. \tag{6.46}$$

Then, n_1 may have two values, namely $n_1 = n$ for even-order denominators and $n_1 = n - 1$ for odd-order ones. This gives two approximate equations $2\pi/\tau_1 \approx 2n$ and $2\pi/\tau_1 \approx 2n_1 + 1$ for the even- and odd-part superscripts of the denominator.

Finally, to obtain the transfer function for the input step voltage, one has to multiply the approximating ratio $U_a(s)$ by Ks. This finally gives

$$H(s) = \frac{KA_1 s \prod_{l=1}^{\nu} \left[s^2 + (2l/\xi)^2\right]}{\prod_{k=0}^{n_1} \left[s^2 + (2k+1)^2\right] + A_2 s \prod_{k=1}^{n} (s^2 + 4k^2)}. \tag{6.47}$$

Representing this $H(s) = N(s)/[D_2(s) + A_2 s D_1(s)]$ as in Equation 6.13, one obtains $y_{21}(s) = -N(s)/D_2(s)$ and $y_{22}(s) = A_2 s D_1(s)/D_2(s)$, two parameters that are sufficient to realize the network within a constant multiplier for y_{21}.

The maximum value of K is defined by the Fialkov condition [9], which specifies that for the unbalanced two ports, the transfer function numerator coefficients should be less or equal to the denominator coefficients for the corresponding degrees of s. The finalized voltage source efficiency coefficient depends on the realized value of A_1.

As an example, we consider the network with the step response that approximates the delayed quasi-rectangular pulse with the following parameters: the delay time $\tau_a = \pi/6$, the duration of each front is $\tau_1 = \pi/6$, the pulse duration calculated at the level of 0.5 of the output amplitude is $\tau_0 = \pi/2$, and the total pulse duration is $\tau = (2\pi)/3$. The reader may notice that this pulse was used in Figure 6.4 to illustrate the problem. The dashed line shows the pulse shape corresponding to the initial assumptions described by Equation 6.28.

Using the pulse parameters one can find that $\xi = (\tau - \tau_1)/\pi = 0.5$. Then, using Equations 6.41 and 6.46, one can find that $\nu = 3$, $n = 6$, and $n_1 = 6$. This allows one to find the ratio that approximates the Laplace transform (Equation 6.36) as

$$U_a(s) = \frac{A_1(s^2 + 16)(s^2 + 64)(s^2 + 144)}{\prod_{k=0}^{6} \left[s^2 + (2k+1)^2\right] + A_2 s \prod_{k=1}^{6} (s^2 + 4k^2)}. \tag{6.48}$$

Equating Equations 6.36 and 6.48 at $s = 0$ and $s = j$, one can find $A_1 = 1.945 \times 10^5$ and $A_2 = 1.329$. The original, $u_a(t)$, corresponding to this transform is shown in Figure 6.5. The network transfer function, thus, is given by the ratio

$$H(s) = \frac{K \times 1.945 \times 10^5 s(s^2 + 16)(s^2 + 64)(s^2 + 144)}{\prod_{k=0}^{6} \left[s^2 + (2k+1)^2\right] + 1.329 s \prod_{k=1}^{6} (s^2 + 4k^2)}. \tag{6.49}$$

Finally, one divides the numerator and denominator of $H(s)$ by $\prod_{k=0}^{6} \left[s^2 + (2k+1)^2\right]$ and obtains

$$y_{22} = \frac{1.329 s \prod_{k=1}^{6} (s^2 + 4k^2)}{\prod_{k=0}^{6} \left[s^2 + (2k+1)^2\right]} \tag{6.50}$$

and

$$y_{21} = \frac{K \times 1.945 \times 10^5 s(s^2 + 16)(s^2 + 64)(s^2 + 144)}{\prod_{k=0}^{6} \left[s^2 + (2k+1)^2\right]}. \tag{6.51}$$

Using the Fialkov condition, one can find that the maximal achievable voltage source efficiency coefficient is $K = 0.984$.

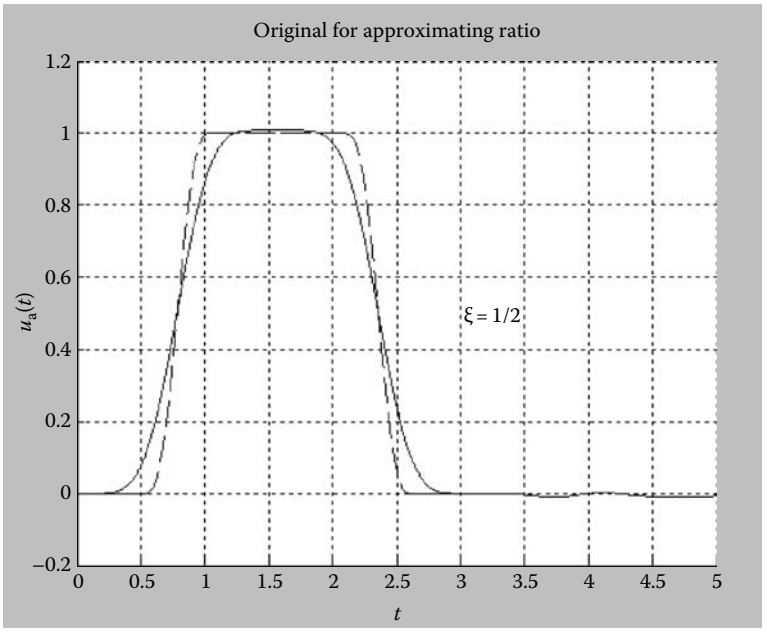


FIGURE 6.5 Step response of the pulse-shaping network.

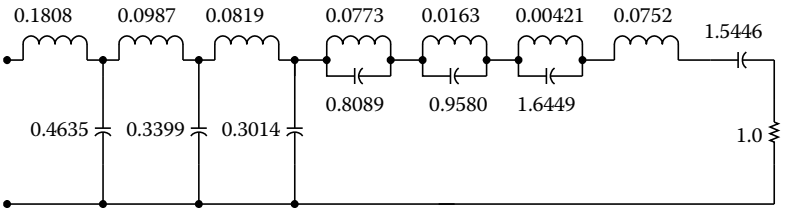


FIGURE 6.6 Realization of the pulse-shaping network.

A realization of this transfer function is shown in Figure 6.6. It starts by realization of common zeros of y_{22} and y_{21} (i.e., by realization of poles of y_{22}^{-1} at $s=0$, $s\to\infty$, $s=\pm j4$, $s=\pm j8$, and $s=\pm j12$) and finishes by the Cauer form to realize other zeros of y_{21} located at $s\to\infty$.

One can verify that the realized value of A_1 is equal to 1.911×10^5. Hence, the realized value of K will be less than the maximum achievable one and equal to 0.967.

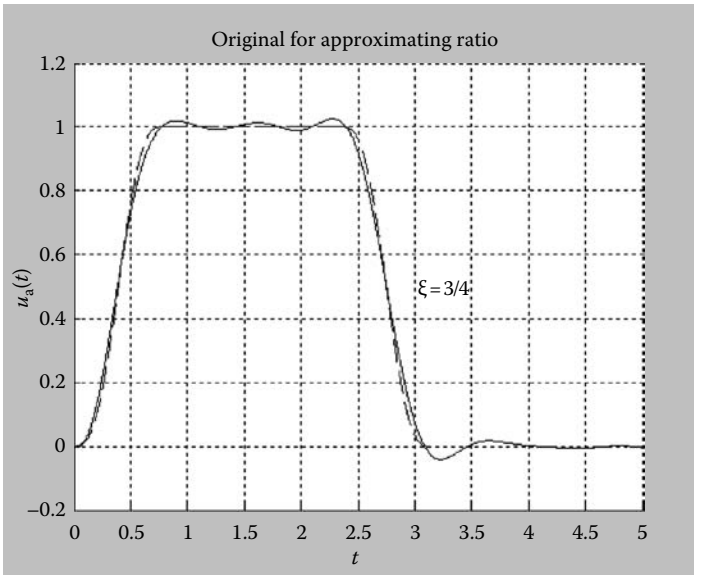
The normalized duration of pulse was $\pi/2$, the normalized value of the load was unit. If the required duration of pulse is τ_r s and the load resistor is R ohms, then each value of inductance in the network of Figure 6.6 should be multiplied by $2\tau_r R/\pi$ and the value of each capacitor by $2\tau_r/(R\pi)$.

Table 6.1 shows that with increase of delay, the maximum voltage source efficiency coefficient increases as well so that the output pulse amplitude can be larger than the amplitude of the input step voltage. The table was calculated for the pulses with the same ratio $\tau_1/\tau_0=1/3$ when the shape of pulse in absolute time is preserved, and in all cases $\nu=3$. We also indicated the required transfer function parameters as well.

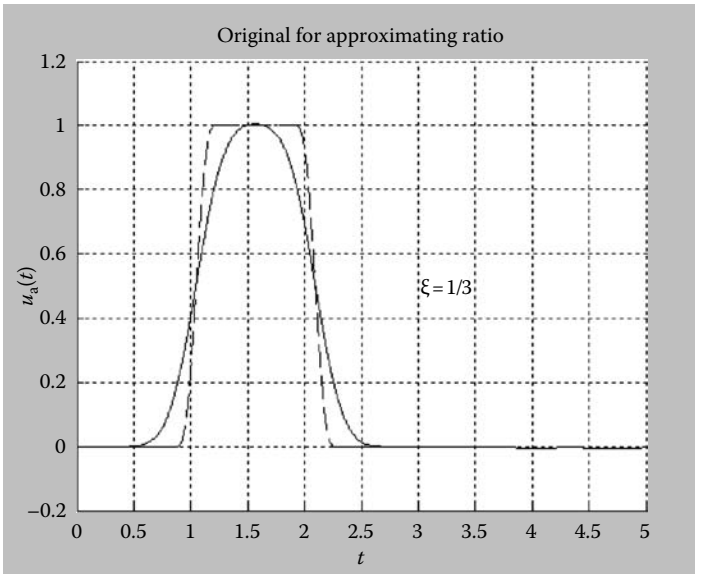
To better visualize the deflection of the realized shapes of the output pulses from the initially assumed shapes, Figure 6.7a and b shows the output pulses for the first and the third lines, respectively, of the

TABLE 6.1 Delay Influence on Voltage Source Efficiency

$\xi = \tau_0/\pi$	τ_1/π	τ_a/π	K	$n = n_1$	A_1	A_2
3/4	1/4	0	0.663	4	1.625×10^{2}	9.462
1/2	1/6	1/6	0.984	6	1.945×10^{5}	13.292
1/3	1/9	5/18	1.476	9	2.673×10^{11}	19.886



(a)



(b)

FIGURE 6.7 Deterioration of output pulse with increasing delay: (a) $\xi = 3/4$ (no delay) and (b) $\xi = 1/3$, $\tau_a/\pi = 5/18$.

Table 6.1 as well. One can see that with increasing delay, the slew rate of the fronts deteriorates and differs from that initially assumed. The duration of the pulse at the level of 0.5 is preserved.

If the zeros of y_{12} coincide with the zeros of y_{22}, the realization is simplified. This was the case in the considered example (it is easy to verify that it is also valid for the transfer function obtained from the third line of Table 6.1). Usually, it can be easily done for one zero only.

6.3 Transfer Functions of Wideband Amplifiers

Wideband amplifiers with a monotonic step response are frequently used for amplification of pulse signals. Their investigation was started by Elmore [13], and the basic results of this work are still used in the design [14–16]. Elmore's approach allowed one to obtain the relationships between the delay time, rise time, and the number of stages for such amplifiers. Yet, this approach introduces a strong limitation: it requires that the poles of transfer functions should be located on the negative real axis of the *s*-plane [13]. The synthesis of the filters with maximally flat delay (Bessel filters) [17] and the Gaussian-response filters [18] removed this restriction. But the transfer functions of these filters were found indirectly, from the requirements in the frequency domain. There is no immediate relationship between the delay-to-rise-time ratio (which is the main parameter in the design of the amplifiers with monotonic step response) and the filter order. One has to simulate their step responses using the available tables of transfer function poles [19], evaluate the delay-to-rise-time ratio, and then decide upon the required order of the transfer function.

It was shown in [4] how to find the transfer function numerically with a monotonic step response starting from the requirements to the delay-to-rise-time ratio. Later on, this problem was solved in [20], yet the results of this work are not easy to apply for design of wideband amplifier transfer functions. Finally, follow the approach proposed in [5], it was shown in [21,22] how to find the transfer functions with monotonic step response and optimized delay-to-rise-time ratio.

Here, we consider a particular case of the solution given in [21]. The amplifier transfer functions tabulated below are obtained when the amplifier impulse response is approximated by the period of $\sin^2(\pi\, t/\tau)$ function (here $\tau \le \pi$). This period is symmetrically located within the interval $0 \le t \le \pi$ of the normalized time t. Then, one repeats the above outlined approximation steps. One finds the Laplace transform of the output response. Approximating further this Laplace transform by a suitable transfer function, one can find an all-pole realizable transfer function of ''maximal for chosen τ order.''

The results of this synthesis procedure are represented by the table of step-response parameters for the transfer functions from the fourth to tenth order. The parameters are also compared with Bessel and Gaussian filter transfer functions of the corresponding order. The table of corresponding transfer functions poles is also given. We do not give the realization of the tabulated transfer functions by LC two-port networks loaded by resistor. The transfer functions of wideband amplifiers are usually realized by active networks, and their realization can be found elsewhere [23]. The proposed method is easily extended on the synthesis of delay networks [24].

6.3.1 Parameters of the Step Response and Its Laplace Transform

Here, the amplifier impulse response, $h(t)$, is approximated by the period of the $\sin^2(\pi\, t/\tau)$ function (here $\tau \le \pi$) symmetrically located within the interval $0 \le t \le \pi$ of the normalized time t, i.e.,

$$h(t) = \begin{cases} 0 & 0 \le t \le (\pi - \tau)/2 \\ A \sin^2 \frac{\pi}{\tau}\left(t - \frac{\pi}{2} + \frac{\tau}{2}\right) & \frac{\pi-\tau}{2} \le t \le \frac{\pi+\tau}{2} \\ 0 & (\pi + \tau)/2 \le t \le \pi \end{cases}. \tag{6.52}$$

The amplitude, A, of the impulse response will be chosen from the normalization condition of the step response, $u(t)$, so that

$$\int_{(\pi-\tau)/2}^{(\pi+\tau)/2} h(t)\mathrm{d}t = 1. \tag{6.53}$$

One finds that $A = 2/\tau$. Then, the delay time $\tau_d = \pi/2$, the rise time $\tau_r = 1/A$, and the delay-to-rise-time ratio, ρ, is equal to

$$\rho = \tau_d/\tau_r = \pi/\tau. \tag{6.54}$$

Using the tables [7] and the delay theorem [8], one finds the Laplace transform of Equation 6.52 as

$$F(s) = \frac{4(\pi/\tau)^2(1-\mathrm{e}^{-\tau s})}{\tau\, s\left[s^2+(2\pi/\tau)^2\right]}\mathrm{e}^{-\left(\frac{\pi-\tau}{2}\right)s}. \tag{6.55}$$

In the following, we will also need the numerical values of this transform at $s = 0$ and $s = j$. One can find that $F(0) = 1$ and $F(j) = -j\frac{8(\pi/\tau)^2\sin(\tau/2)}{\tau\left[(2\pi/\tau)^2-1\right]}$.

6.3.2 Transfer Function Approximation

The Laplace transform (Equation 6.55) can be rewritten as

$$F(s) = \frac{8(\pi/\tau)^2\sinh(\tau s/2)}{\tau s\left[s^2+(2\pi/\tau)^2\right]}\mathrm{e}^{-s\frac{\pi}{2}}. \tag{6.56}$$

Using the expansion (Equation 6.37) for the function of sinh x, and approximating the infinite product with two terms, as

$$\sinh(s\tau/2) \approx \frac{s\tau}{2}\left(1+\frac{s^2\tau^2}{4\pi^2}\right), \tag{6.57}$$

one finds that $F(s)$ can be approximated as

$$F(s) \approx \mathrm{e}^{-s\frac{\pi}{2}}. \tag{6.58}$$

The validity region of this approximation is determined by the first term in the discarded part of the approximation (Equation 6.57). On the $j\omega$-axis, the border for the maximal frequency, ω_m, where the approximation (Equation 6.58) is still valid, is thus, determined by the inequalities

$$(2\pi)/\tau < \omega_m < (4\pi)/\tau. \tag{6.59}$$

On the $j\omega$-axis, the function (Equation 6.58) becomes

$$F_a(j\omega) = \cos(\omega\pi/2) - j\sin(\omega\pi/2). \tag{6.60}$$

The functions $\cos(\omega\pi/2)$ and $\sin(\omega\pi/2)$ can also be represented as infinite products as in Equations 6.17 and 6.18, one can write that

$$F_a(j\omega) = \prod_{i=0}^{\infty}\left(1 - \frac{\omega^2}{(2i+1)^2}\right) - j\frac{\omega\pi}{2}\prod_{i=1}^{\infty}\left(1 - \frac{\omega^2}{4i^2}\right). \tag{6.61}$$

Using the finite number of terms in the products, one can approximate $F(j\omega)$ as

$$F_a(j\omega) \approx \prod_{i=0}^{n_1}\left(1 - \frac{\omega^2}{(2i+1)^2}\right) - j\frac{\omega\pi}{2}\prod_{i=1}^{n}\left(1 - \frac{\omega^2}{4i^2}\right). \tag{6.62}$$

This result allows one to find a realizable transfer function

$$H_a(s) = \frac{A_1}{\prod_{i=0}^{n_1}\left[s^2 + (2i+1)^2\right] + A_2 s\prod_{i=1}^{n}(s^2 + 4i^2)} \tag{6.63}$$

that approximates $F(s)$ given by Equation 6.56.

The following two possibilities should be considered for n_1 and n. If $n_1 = n$, then the denominator degree is even, if $n_1 = n - 1$ then the denominator degree is odd. Indeed, if these conditions are satisfied, and A_1 and A_2 are positive, then the denominator of $H_a(s)$ is a Hurwitz polynomial [9]. The denominator roots will be located in the left half of the s-plane, and this $H_a(s)$ can be realized as a cascade connection of first- and second-order low-pass stages.

The approximation will be the best if one takes the maximum possible value of n_1 or n in Equation 6.63. Substituting $s = j\omega$ in Equation 6.63, one can find that

$$H_a(j\omega) = \frac{A_1\prod_{i=0}^{n_1}\left[(-\omega^2 + (2i+1)^2\right] - jA_2\omega\prod_{i=1}^{n}(-\omega^2 + 4i^2)}{\left(\prod_{i=0}^{n_1}\left[\left(-\omega^2 + (2i+1)^2\right)\right]^2 + A_2^2\omega^2\left(\prod_{i=1}^{n}(-\omega^2 + 4i^2)\right)^2\right)} \tag{6.64}$$

Comparing Equations 6.64 and 6.62, one can see that to have the maximal order of the transfer function one has to take the maximal value of n_1 or n in Equation 6.62 as well. But the choice of n_1 and n is defined, as it follows from Equation 6.59, by the following inequalities

$$\begin{cases} (2\pi)/\tau < 2n_1 + 1 < (4\pi)/\tau \\ (2\pi)/\tau < 2n < (4\pi)/\tau \end{cases}. \tag{6.65}$$

One has to choose the largest value satisfying one of these inequalities. If this largest value is $n_1 = m$, then one takes $n = m$ as well and obtains the transfer function of even order. If this largest value is $n = m$, then one takes $n_1 = m - 1$, and obtains the transfer function of odd order.

Finally, to find A_1, one equates $H_a(s)$ and $F(s)$ at $s = 0$. This gives $A_1 = \prod_{i=1}^{n_1}(2i+1)^2$. Equating $H_a(s)$ and $F(s)$ at $s = j$ one finds

$$A_2 = \frac{\tau[(2\pi/\tau) - 1]A_1}{8(\pi/\tau)^2\sin(\tau/2)\prod_{i=1}^{n}\left[(2i)^2 - 1\right]}.$$

6.3.3 Example of Transfer Function Design

Assume that we design an amplifier transfer function with $\rho = 1.6$. This requires that $\tau = 5\pi/8$. Substituting this value in Equation 6.65, one obtains the inequalities

$$\begin{cases} 3\frac{1}{5} < 2n_1 + 1 < 6\frac{2}{5} \\ 3\frac{1}{5} < 2n < 6\frac{2}{5} \end{cases}. \tag{6.66}$$

From these inequalities, one finds that $n = 3$ and $n_1 = 2$. Hence, the realizable transfer function

$$H_a(s) = \frac{A_1}{\prod_{i=0}^{i=2}\left[s^2 + (2i+1)^2\right] + A_2\prod_{i=1}^{i=3}\left[s^2 + (2i)^2\right]} \tag{6.67}$$

of the seventh-order approximating $F_a(s)$ (and, hence $F(s)$ given by Equation 6.56) is found. It has the maximal possible order for the required τ. The constants A_1 and A_2 can be calculated equating $F(s)$ given by Equation 6.56 and $H_a(s)$ given by Equation 6.67 at the points $s = 0$ and $s = j$. One finds that $A_1 = 225$ and $A_2 = 0.1522$. Hence, the realizable transfer function

$$H_a(s) = \frac{225}{(s^2+1)(s^2+9)(s^2+25) + 0.1522s(s^2+4)(s^2+16)(s^2+36)} \tag{6.68}$$

is approximating $F(s)$ for $\tau = 5\pi/8$.

The impulse and step responses for this transfer function are shown in Figure 6.8. One can see that the impulse response of Equation 6.68 is close to the response (Equation 6.52) for $t = 5\pi/8$, and the step

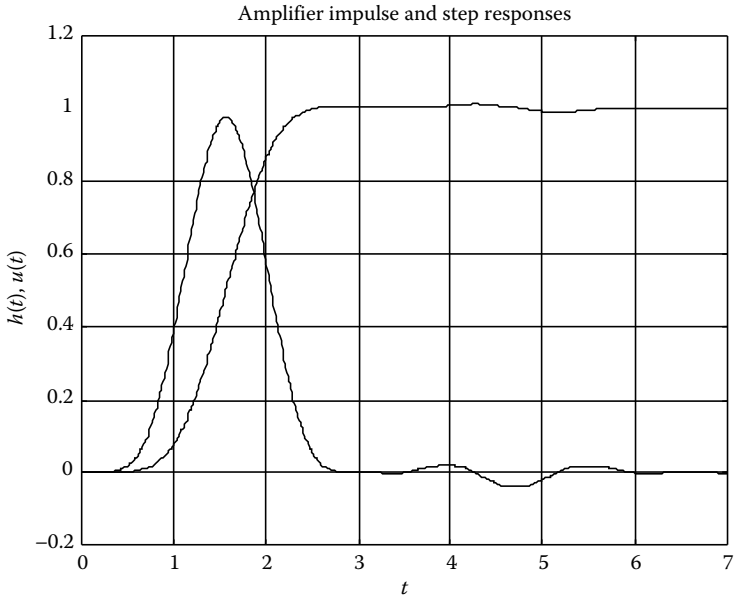


FIGURE 6.8 Example of impulse and step responses.

TABLE 6.2 Step Response Parameters

D	τ	ρ	ρ_s	$Ov(\%)$	$Un(\%)$	ρ_b	$Ov_b(\%)$	ρ_g
4	π	1.00	1.06	1.4	1.9	0.91	0.8	0.87
5	$7\pi/8$	1.15	1.20	0.3	1.5	1.06	0.8	0.98
6	$3\pi/4$	1.33	1.36	1.2	0.8	1.20	0.6	1.11
7	$5\pi/8$	1.60	1.50	1.3	0.9	1.32	0.5	1.21
8	$\pi/2$	2.00	1.63	1.6	0.0	1.42	0.3	1.31
9	$15\pi/32$	2.13	1.74	1.5	0.6	1.52	0.2	1.43
10	$7\pi/16$	2.29	1.86	1.2	0.0	1.63	0.1	1.50

TABLE 6.3 Transfer Functions Poles

D	Poles
4	$-0.664 \pm j2.228$; $-1.103 \pm j0.669$
5	$-0.604 \pm j3.174$; $-1.032 \pm j1.508$; -1.140
6	$-0.532 \pm j4.119$; $-0.949 \pm j2.293$; $-1.259 \pm j0.730$
7	-0.487 ± 5.095; $-0.860 \pm j3.206$; $-1.215 \pm j1.526$; -1.320
8	$-0.460 \pm j6.083$; $-0.806 \pm j4.172$; $-1.148 \pm j2.446$; $-1.230 \pm j0.857$
9	$-0.404 \pm j7.060$; $-0.678 \pm j5.096$; $-0.969 \pm j3.167$; $-1.344 \pm j1.304$; -1.880
10	$-0.400 \pm j8.059$; $-0.677 \pm j6.100$; $-0.980 \pm j4.204$; $-1.272 \pm j2.498$; $-1.312 \pm j0.874$

response is practically monotonic (the overshoot is 1.3% and the undershoot is 0.9%). The realized delay-to-rise-time ratio obtained in simulation is $\rho_s = 1.5$.

6.3.4 Tabulated Results

The results of our derivation are summarized in two tables. Table 6.2 gives the step-response parameters of the derived transfer functions. In Table 6.2, D is the degree of the transfer function denominator, ρ is the estimate of delay-to-rise-time ratio obtained from Equation 6.54, ρ_s is the delay-to-rise-time ratio obtained in simulations of approximating transfer functions, Ov is the step-response overshoot obtained in simulations, and Un is the undershoot also obtained in simulations.

These overshoot and undershoot columns deserve some discussion. Let us return to the example, the step response of the transfer function with $D = 7$ (Figure 6.8). As one can see, the step response $u(t)$ does not have the overshoot as it usually appears in ''normal'' transfer functions, that is, during application of the pulse signal. Here one has a small postpulse wave located in the interval $\pi < t < 2\pi$. One can verify that this is valid for all proposed transfer functions. Finally, Table 6.2 gives the comparison of the proposed transfer functions with the transfer functions of the Bessel (ρ_b, Ov_b) and Gaussian (ρ_g, no overshoot) filter transfer functions. Table 6.3 gives the poles of the transfer functions.

6.4 Forming a Sinusoidal Pulse

Recently, most of the industrial emphasis in ultra-wideband technology (UWB) has focused on the short range, high-data-rate applications. However, due to its low power properties, impulse UWB is also suitable for low-power, low-data-rate applications [25]. One of the challenges in such low-power systems is how to efficiently generate the pulses. Several common pulse waveforms were studied, with the relation between their spectral characteristics and waveform parameters pointed out [26,27]. The approximation method described here allows one to synthesize the pulse-forming reactance networks loaded by resistors, and these networks may represent, in applications to UWB systems, model of transmitting antennas.

Comparing the complexities of the network forming pulses of different shapes gives additional information for UWB system design. The developed synthesis approach is suitable for many pulse forms used in UWB systems, but for the reasons of space limitations, we consider only an example of synthesis of the network forming sinusoidal pulse with sinusoidal envelope [28]. In [29], one can find information on the pulse-forming network for monocycle pulse.

6.4.1 Required Transfer Function

Let the required output pulse $u(t)$ be a sinusoidal oscillation of a radian frequency ν with a sinusoidal envelope, $\sin \Omega t$ that has the finite duration of $\Omega t_d = \pi$ in the normalized time (Figure 6.9), that is,

$$u(t) = \begin{cases} g(t) \sin \nu t = \sin \Omega t \sin \nu t & 0 \le \Omega t \le \pi \\ 0 & \Omega t > \pi \end{cases}. \tag{6.69}$$

It is assumed here that Equation 6.69 is the impulse or a step response. Then using the shift theorem [8], one can find that the Laplace transform of the envelope $g(t)$ is

$$G(s) = \frac{\Omega(1 + e^{-\pi s})}{s^2 + \Omega^2} \tag{6.70}$$

Considering that $\sin \nu t = (e^{j\nu} - e^{-j\nu})/(2j)$ and using the theorem of complex translation [8], one writes that the Laplace transform of $u(t)$ is equal to

$$F(s) = \frac{\Omega}{2j}\left[\frac{1 + e^{-\pi(s-j\nu)}}{(s-j\nu)^2 + \Omega^2} - \frac{1 + e^{-\pi(s+j\nu)}}{(s-j\nu)^2 + \Omega^2}\right]. \tag{6.71}$$

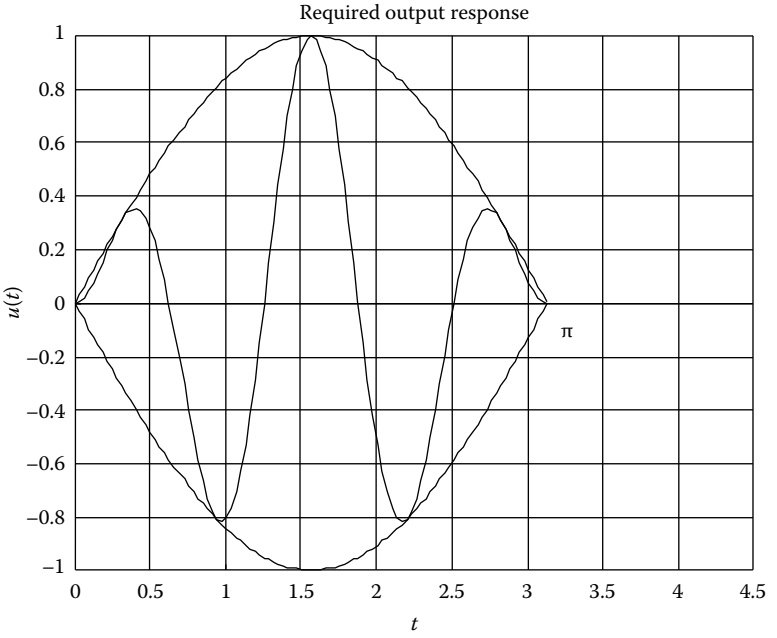


FIGURE 6.9 Required output response with sinusoidal envelope.

After some simple algebra, one obtains that

$$F(s) = \frac{2\nu\Omega s\left[1-(-1)^{\nu+1}e^{-\pi s}\right]}{\left[s^2+(\nu+\Omega)^2\right]\left[s^2+(\nu-\Omega)^2\right]}. \tag{6.72}$$

Considering that $\Omega = 1$ and assuming that ν is an integer, one can rewrite that

$$F(s) = \frac{2\nu s\left[1-(-1)^{\nu+1}e^{-\pi s}\right]}{\left[s^2+(\nu+1)^2\right]\left[s^2+(\nu-1)^2\right]}. \tag{6.73}$$

It is possible to consider two cases. If $\nu+1$ is even, then Equation 6.73 becomes

$$F_{ev}(s) = \frac{4\nu s e^{-\frac{\pi}{2}s}\sinh\left(\frac{\pi}{2}s\right)}{\left[s^2+(\nu+1)^2\right]\left[s^2+(\nu-1)^2\right]}. \tag{6.74}$$

If $\nu+1$ is odd, then Equation 6.73 becomes

$$F_{odd}(s) = \frac{4\nu s e^{-\frac{\pi}{2}s}\cosh\left(\frac{\pi}{2}s\right)}{\left[s^2+(\nu+1)^2\right]\left[s^2+(\nu-1)^2\right]}. \tag{6.75}$$

Then, Equations 6.74 and 6.75 do not belong to a class of realizable transfer functions if they are the transforms of impulse responses. Their multiplication by s (for the case if they are transforms of the step responses) does not give realizable functions either. One has again to truncate the decompositions into products for transcendental functions, and to find the required final approximation.

6.4.2 Approximation for Realization

Now, we find the approximations to the Laplace transforms (Equations 6.74 and 6.75) resulting in the realizable transfer functions. Let us restrict ourselves by the last case ($\nu+1$ is odd) only. The other case can be considered in a similar way. Substituting $s = j\omega$, one can find that

$$\mathrm{Re}F_{odd}(j\omega) = \frac{2\nu\omega\sin\left(\frac{\pi}{2}\omega\right)\cos\left(\frac{\pi}{2}\omega\right)}{\left[-\omega^2+(\nu+1)^2\right]\left[-\omega^2+(\nu-1)^2\right]} \tag{6.76}$$

and

$$\mathrm{Im}F_{odd}(j\omega) = \frac{2\nu\omega\cos^2\left(\frac{\pi}{2}\omega\right)}{\left[-\omega^2+(\nu+1)^2\right]\left[-\omega^2+(\nu-1)^2\right]}. \tag{6.77}$$

Taking Equations 6.17 and 6.18 into consideration, one can see that $\mathrm{Re}F_{odd}(j\omega)$ and $\mathrm{Im}F_{odd}(j\omega)$ have common zeros at $\omega = 0$; and $\omega = \pm1$; ±3; ±5; … (due to the common multiplier $\cos(\pi\omega/2)$), yet the zeros $\omega = \pm(\nu-1)$ and $\omega = \pm(\nu+1)$ of this multiplier are not common zeros (they are canceled by the poles). Then, the $\mathrm{Re}F_{odd}(j\omega)$ has the zeros at $\omega = 0$; ±2; ±4; ±6; … (due to the multiplier $\sin(\pi\omega/2)$), and the $\mathrm{Im}F_{odd}(j\omega)$ has the zeros at $\omega = \pm1$; ±3; ±5; … (due to the second multiplier of $\cos(\pi\omega/2)$). Hence in the last two groups, the zeros are alternating. One can see that the ratio $H_{odd}(s)$ for approximation of $F(s)$ should have the form

$$H_{\text{odd}}(s) = \frac{A_1 s \prod_{k=0,\, k\neq\nu-1,\, k\neq\nu+1}^{k=\rho} \left[s^2 + (2k+1)^2\right]}{\prod_{k=0}^{n} \left[s^2 + (2k+1)^2\right] + A_2 s \prod_{k=1}^{n_1} (s^2 + 4k^2)}. \tag{6.78}$$

Indeed, if one calculates $\text{Re}H_{\text{odd}}(j\omega)$ and $\text{Im}H_{\text{odd}}(j\omega)$ and compare them with $\text{Re}F_{\text{odd}}(j\omega)$ and $\text{Im}F_{\text{odd}}(j\omega)$, one finds that the zeros of $\text{Re}H_{\text{odd}}(j\omega)$ and $\text{Im}H_{\text{odd}}(j\omega)$ have the same properties: there are common zeros (due to the numerator) and alternating zeros (due to even and odd polynomials in the denominator). The number of zeros is, of course, finite, and is defined by the order of $H_{\text{odd}}(s)$. In practice it is sufficient to choose $\rho = \nu + 2$. Then one chooses $n = n_1$ for even and $n = n_1 - 1$ for odd polynomials in the denominator. The maximal of these two numbers $\max(n, n_1) \geq \rho$. The approximation becomes better when the order of $H_{\text{odd}}(s)$ increases. The constants A_1 and A_2 can be found equating $F_{\text{odd}}(s)$ and $H_{\text{odd}}(s)$ at two points of the s-plane. The case of $F_{\text{ev}}(s)$ is treated in a similar way.

6.4.3 Example

As an example, we consider the network that should have the impulse response

$$u(t) = \begin{cases} \sin t \sin 5t & 0 \leq t \leq \pi \\ 0 & t > \pi \end{cases}. \tag{6.79}$$

(It is this pulse that was shown in Figure 6.9.) We have $\nu + 1 = 6$, that is, the even case. It is easy to find that the Laplace transform of Equation 6.79 is equal to

$$F_{\text{ev}}(s) = \frac{20 s e^{-\frac{\pi}{2}s} \sinh\left(\frac{\pi}{2}s\right)}{(s^2 + 4^2)(s^2 + 6^2)}. \tag{6.80}$$

The real and imaginary parts of this function are

$$\text{Re}F_{\text{ev}}(j\omega) = -\frac{20\omega \cos\left(\frac{\pi}{2}\omega\right) \sin\left(\frac{\pi}{2}\omega\right)}{(-\omega^2 + 16)(-\omega^2 + 36)} \tag{6.81}$$

and

$$\text{Im}F_{\text{ev}}(j\omega) = \frac{20\omega \sin^2\left(\frac{\pi}{2}\omega\right)}{(-\omega^2 + 16)(-\omega^2 + 36)}. \tag{6.82}$$

In the interval $0 \leq \omega \leq 8$, the functions $\text{Re}F_{\text{ev}}(j\omega)$ and $\text{Im}F_{\text{ev}}(j\omega)$ have double common zero at $\omega = 0$, and simple common zeros at $\omega = \pm 2$ and $\omega = \pm 8$. The zeros at $\omega = \pm 4$ and $\omega = \pm 6$ are not common zeros, after cancelation of poles at $\omega = \pm 4$ and $\omega = \pm 6$ they disappear from $\text{Re}F_{\text{ev}}(j\omega)$ and left in $\text{Im}F_{\text{ev}}(j\omega)$ only. The function $\text{Re}F_{\text{ev}}(j\omega)$ has zeros at $\omega = \pm 1; \pm 3; \pm 5; \pm 7$. These zeros are alternating with the zeros $\omega = 0\ \pm 2; \pm 4; \pm 6; \pm 8$ that occur in $\text{Im}F_{\text{ev}}(j\omega)$ due to the second $\sin(\pi\omega/2)$ multiplier in the numerator of Equation 6.82. As a result, one can use the ratio

$$H_{\text{ev}}(s) = \frac{A_1 s^2 (s^2 + 4)(s^2 + 64)}{[(s^2 + 1)(s^2 + 9)(s^2 + 25)(s^2 + 49) + A_2 s (s^2 + 4)(s^2 + 16)(s^2 + 36)(s^2 + 64)]}. \tag{6.83}$$

The coefficients A_1 and A_2 can be found, for example, equating Equations 6.81 and 6.84 at $s = j4$ and $s = j5$. One finds that $A_1 = 2.65762$ and $A_2 = 0.13288$.

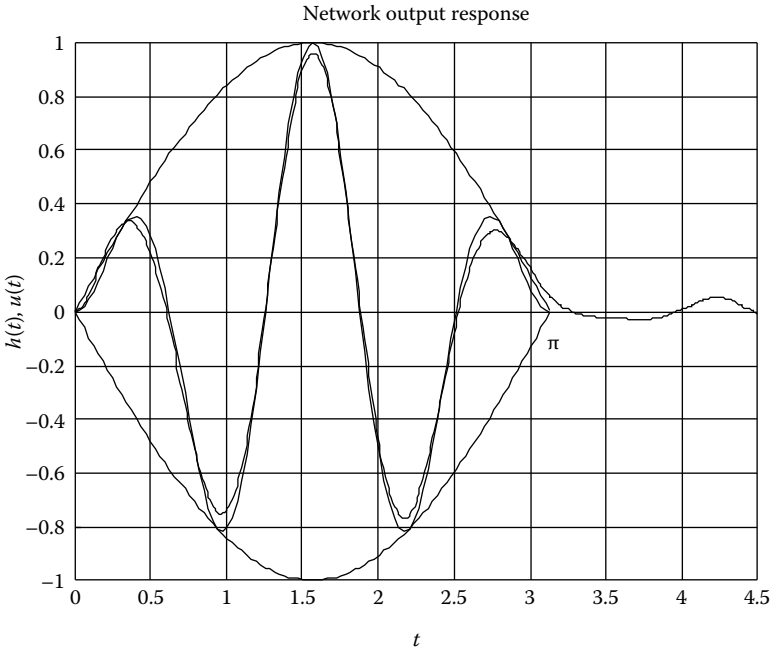


FIGURE 6.10 Network output response.

Because Equation 6.80 is the impulse response, one takes directly

$$H(s) = H_{ev}(s). \tag{6.84}$$

Figure 6.10 shows the inverse Laplace transform $h(t) = \mathbf{L}^{-1}(H(s))$ of this transfer function that is compared with $u(t)$ given by Equation 6.79.

Dividing the numerator and denominator of $H(s) = H_{ev}(s)$ by the odd part of the denominator one writes that

$$y_{21} = -\frac{A_1 s}{A_2(s^2+16)(s^2+36)} \tag{6.85}$$

and

$$y_{22} = \frac{(s^2+1)(s^2+9)(s^2+25)(s^2+49)}{A_2 s(s^2+4)(s^2+16)(s^2+36)(s^2+64)}. \tag{6.86}$$

The LC-realization using these two parameters is shown in Figure 6.11. The realization is done within a constant multiplier for y_{21}. It starts by realization of the private poles of y_{22} at $s = 0$, $s = \pm 2$, and $s = \pm j8$. The residual admittance

$$y'_{22} = \frac{2.850s(s^2+24.936)}{(s^2+16)(s^2+36)} \tag{6.87}$$

is realized, first, by subtraction of the poles of $1/(y'_{22})$ at $s = 0$ and $s = \infty$. This provides the zeros at $s = 0$ and $s = \infty$ common for y_{21} and y'_{22}. Then, after subtraction of the series elements corresponding to these

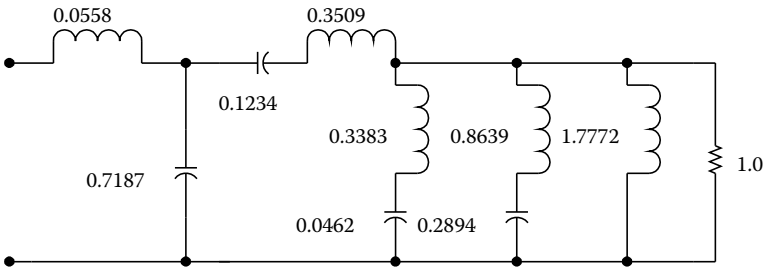


FIGURE 6.11 Realization of the pulse-shaping network.

poles of $1/(y'_{22})$, the final residual conductance is realized to provide two additional zeros of y_{21} at $s = \infty$. One can find that this realization provides the level of y_{21} that is 3.553 times higher than that given by Equation 6.85.

One can see that multiplication of Equation 6.83 by s gives the realizable function as well. The realization of the network providing Equation 6.79 as the step response is not given here for the reason of space limitation.

6.5 Summary

The synthesis of a pulse-forming network with output response that is close to a rectangular shape pulse requires two step approximation procedure. The derivative of this output response is described by positive and delayed negative semi-periods of the sine-squared function. This is the first step of approximation procedure. The real and imaginary parts of Laplace transform of thus approximated output pulse are expanded in infinite products. Then, using a finite number of terms in these products, one obtains an algebraic ratio that approximates this Laplace transform. This is the second step of approximation procedure.

If the input excitation is the unit impulse function, the obtained ratio is directly the required transfer function realizable by a reactance network loaded by resistor. Two closely connected realizable transfer functions (providing the same output) are obtained by simple multiplication of the previously obtained transfer function by s (if the input excitation is the unit step function) or by $(s^2 + \Omega^2)$ (if the input excitation is the sinusoid of unit amplitude and of frequency Ω).

The synthesis of a linear network shaping a nonperiodic (in the considered case quasi-rectangular) pulse from the sinusoidal voltage is using the short period of time after turning on this input voltage. The zeros of transfer function are used to reject the periodic solution and create zero solution during the network steady-state operation. The efficiency of this pulse-shaping circuit may be insufficient; yet, conceptually the synthesis procedure is not very much different from the synthesis of pulse-shaping networks using for this purpose the discharge of a capacitor.

The two-step approximation resulting in the realizable reactance networks may be extended on the wide variety of output pulses including wideband amplifier transient responses and sinusoidal pulses of finite duration and with wide variety of envelop shape.

The method is simple, and the output response is approximated with a small error. The network complexity to obtain better approximations can be easily controlled. The approximation precision as well as the network complexity can be improved using computer methods [30].

The proposed procedure for synthesis of pulse-forming networks is based on approximation of the meaningful part of the output signal spectrum. The zeros of real and imaginary parts of approximated and approximating functions coincide in the meaningful part of the spectrum, and their amplitude values are close to each other. This closeness of spectrums provides good approximation in the time domain.

References

1. G. N. Glasoe, J. V. Lebacqz, eds., The pulse-forming network, in *Pulse Generators*, Chapter 6, Boston Technical Publishers, Boston, MA, pp. 175–224, 1964.
2. P. W. Smith, *Transient Electronics*, Wiley, Chichester, England, 2002.
3. B. E. Thomson, The synthesis of a network to have a sine-squared impulse response, *IEE Proceedings*, 99, Pt. III, 373–376, 1952.
4. J. Jess, H. W. Schussler, On design of pulse-forming networks, *IEEE Transactions on Circuit Theory*, CT-12, 393–400, 1965.
5. P. N. Matkhanov, *Time-Domain Synthesis of Reactance Two-Ports*, Energia Publishing House, Leningrad, USSR, 1970 (in Russian).
6. I. M. Filanovsky, P. N. Matkhanov, Synthesis of reactance networks shaping a quasi-rectangular pulse, *IEEE Transactions on Circuits and Systems*, Part II, 52, 242–245, 2005.
7. V. A. Ditkin, A. P. Prudnikov, *Integral Transforms and Operational Calculus*, Pergamon Press, Oxford, 1965.
8. R. V. Churchill, *Operational Mathematics*, 2nd edn., McGraw-Hill, New York, 1958.
9. N. Balabanian, *Network Synthesis*, Prentice-Hall, Englewood Cliffs, NJ, 1958.
10. L. Weinberg, *Network Analysis and Synthesis*, McGraw-Hill, New York, 1962.
11. P. J. Nahin, *An Imaginary Tale*, Chapter 6, Princeton University Press, Princeton, NJ, 1998.
12. I. M. Filanovsky, P. N. Matkhanov, Synthesis of a pulse-forming reactance network shaping a quasi-rectangular delayed output pulse, *IEEE Transactions on Circuits and Systems*, Part II, 51, 190–194, 2004.
13. W. C. Elmore, The transient response of damped linear networks with particular regard to wideband amplifiers, *Journal of Applied Physics*, 19, 55–63, 1948.
14. D. L. Feucht, *Handbook of Analog Circuit Design*, Chapter 8, Academic Press, Inc., San Diego, CA, pp. 318–379, 1990.
15. T. H. Lee, Risetime, delay and bandwidth, in *The Design of CMOS Radio-Frequency Integrated Circuits*, Chapter 7, Cambridge University Press, New York, pp. 167–174, 1998.
16. P. Staric, E. Margan, *Wideband Amplifiers*, Springer, Dordrecht, The Netherlands, 2006.
17. L. Storch, Synthesis of constant time delay ladder networks using Bessel polynomials, *Proceedings of IRE*, 42, 1666–1675, 1954.
18. M. Dishal, Gaussian-response filter design, *Electrical Communication*, 36, 3–26, 1959.
19. C. S. Lindquist, *Active Network Design*, Steward & Sons, Long Beach, CA, 1977.
20. G. C. Temes, The prolate filter: An ideal low-pass filter with optimum step-response, *Journal of Franklin Institute*, 293(2), 77–103, 1972.
21. I. M. Filanovsky, A new class of wide-band amplifiers with monotonic step response, *IEEE Transactions on Circuits and Systems*, Part I, 50, 569–571, 2003.
22. I. M. Filanovsky, P. N. Matkhanov, On the synthesis of wideband multistage amplifiers with monotonic step response, *Circuits, Systems and Signal Processing*, 21(6), 567–580, 2002.
23. J. Davidse, *Analog Electronic Circuit Design*, Prentice Hall, Englewood Cliffs, NJ, 1991.
24. I. M. Filanovsky, P. N. Matkhanov, Synthesis of time delay networks approximating the pulse response described by an integer power of a sinusoid over its semiperiod, *Analog Integrated Circuits and Signal Processing*, 28, 81–88, 2001.
25. S. A. P. Haddad, Ultra low-power biomedical signal processing, PhD thesis, Delft University of Technology, The Netherlands, 2006.
26. X. Chen, S. Kiaei, Monocycle shapes for ultra wideband system, *IEEE Proceedings of the International Symposium on Circuits and Systems, ISCAS'2002*, 1, 597–600, 2002.
27. B. Parr, ByungLok Cho, K. Wallace, Zhi Ding, A novel ultra-wideband pulse design algorithm, *IEEE Communication Letters*, 7, 219–221, 2003.

28. I. M. Filanovsky, P. N. Matkhanov, On synthesis of a reactance network having the step response described by a sinusoid with a given envelope, *IEEE Proceedings of the International Symposium on Circuits and Systems, ISCAS'2001*, 1, 603–606, 2001.
29. I. M. Filanovsky, X. Dong, On synthesis of reactance network shaping a monocycle pulse, *IEEE Proceedings of the 47th Midwest Symposium on Circuits and Systems,* 1, 101–104, 2004.
30. M. Vucic, G. Molnar, Design of systems with prescribed impulse response based on second-order cone programming, *IEEE Proceedings of the International Symposium on Circuits and Systems, ISCAS'2007*, 3311–3313, 2007.

7

Fundamentals of Digital Signal Processing

7.1 Introduction 7-1
Fourier Series for Continuous-Time Periodic Signals • Example and Discussion • Discrete-Time Signals • Discrete Fourier Transform • Examples and Discussion • DFT Applications • Fast Fourier Transform • Continuous-Time Fourier Transform • Examples and Discussion • Discrete-Time Fourier Transform • DTFT Model • Sampling Theorem • DTFT Properties • Example and Discussion • Linear and Time Invariant Discrete-Time Systems • Convolution • Stability • Frequency Response • Examples and Discussion • Ideal Digital Filters • z-Transform • Bilateral z-Transform Properties • z-Plane • Transfer Function • Unilateral z-Transform • Conclusion
References 7-36

Roland Priemer
University of Illinois at Chicago

7.1 Introduction

Real-time and off-line processing of continuous-time signals by digital means (digital signal processing [DSP]) has become a viable processing mode over analog means for several reasons, some of which are digital signal processors, microcontrollers, and microprocessors. These are inexpensive, programmable, reproducible, consume low power, have computing speeds suitable for signals with bandwidths beyond base band video, and can operate in extreme environments. Some broad application areas of DSP are automotive industry, consumer electronics, communication systems, and medical systems.

Since real-world signals are continuous in time, the technologist must properly interpret results of processing signals by digital means. This requires an understanding of the origin of and relationship among the basic tools used for DSP. We shall first consider the Fourier series (FS) concept for continuous-time periodic signals. Everything that follows will be based on this concept.

This chapter is intended for those who have some experience with the material generally covered in a first course on continuous-time signals and systems, and would like a brief introduction to the fundamentals of DSP.

7.1.1 Fourier Series for Continuous-Time Periodic Signals

Given is a real and periodic signal $x(t)$, which satisfies

$$x(t) = x(t + T_0)$$

for some period T_0 and all t. One period of $x(t)$ is $x_p(t) = x(t)$, for $t_0 \le t < t_0 + T_0$ and $x_p(t) = 0$ for $t < t_0$ and $t \ge t_0 + T_0$, where t_0 is arbitrary. The periodic signal $x(t)$ can be written as the periodic extension of $x_p(t)$, which is

$$x(t) = \sum_{r=-\infty}^{+\infty} x_p(t - rT_0)$$

Let us approximate $x(t)$ by a sum of sinusoidal functions given by

$$\hat{x}(t) = \hat{a}_0 + \sum_{k=1}^{\infty} \hat{a}_k \cos(k\omega_0 t) + \sum_{k=1}^{\infty} \hat{b}_k \sin(k\omega_0 t)$$

where ω_0, called the fundamental frequency, is found with $\omega_0 = 2\pi/T_0$ rad/s and $f_0 = 1/T_0$ Hz. The frequency ω of each sinusoidal function is $\omega = k\omega_0$. Therefore, the approximation $\hat{x}(t)$ is also periodic with period T_0. The approximation error is

$$e(t) = x(t) - \hat{x}(t)$$

We choose the coefficients $\hat{a}_0$, $\hat{a}_k$, and $\hat{b}_k$ to minimize the mean square error given by

$$\varepsilon^2\left(\hat{a}_0, \hat{a}_k, \hat{b}_k\right) = \frac{1}{T_0} \int_0^{T_0} e^2(t)\mathrm{d}t$$

which is a quadratic function of $\hat{a}_0$, $\hat{a}_k$, and $\hat{b}_k$. Denote the $\hat{a}_0$, $\hat{a}_k$, and $\hat{b}_k$ that minimize ε^2 by a_0, a_k, and b_k, respectively. Setting the partial derivatives of ε^2 with respect to $\hat{a}_0$, $\hat{a}_k$, and $\hat{b}_k$ to zero gives

$$a_0 = \frac{1}{T_0} \int_0^{T_0} x(t)\mathrm{d}t$$

which is the average value of $x(t)$, and

$$a_k = \frac{2}{T_0} \int_0^{T_0} x(t) \cos(k\omega_0 t)\mathrm{d}t$$

$$b_k = \frac{2}{T_0} \int_0^{T_0} x(t) \sin(k\omega_0 t)\mathrm{d}t$$

If $x(t)$ satisfies the Dirichlet conditions, which are that $x(t)$ must

1. Have a finite number of extrema in any given time interval
2. Have a finite number of discontinuities in any given time interval
3. Be absolutely integrable over a period

then setting $\hat{a}_0 = a_0$, $\hat{a}_k = a_k$, and $\hat{b}_k = b_k$ gives $\varepsilon^2(a_0, a_k, b_k) = 0$, and we write

$$x(t) = a_0 + \sum_{k=1}^{\infty} a_k \cos(k\omega_0 t) + \sum_{k=1}^{\infty} b_k \sin(k\omega_0 t) \tag{7.1}$$

which is called the FS representation of $x(t)$, and the a_0, a_k, and b_k are the trigonometric FS coefficients of $x(t)$. If $x(t)$ is discontinuous at $t = t_d$, then for $t = t_d$, the FS converges to the average value of $x(t)$ about $t = t_d$, which is

$$\frac{x(t_d^-) + x(t_d^+)}{2} \tag{7.2}$$

Furthermore, if $x(t)$ is everywhere finite and continuous, then we have the stronger result that $e(t) = 0$.

For further development, it is more convenient to apply Euler's identity, which is

$$e^{j\alpha} = \cos(\alpha) + j\sin(\alpha) \tag{7.3}$$

for any real number α, and then $x(t)$ becomes

$$x(t) = a_0 + \sum_{k=1}^{\infty} a_k \frac{e^{jk\omega_0 t} + e^{-jk\omega_0 t}}{2} + \sum_{k=1}^{\infty} b_k \frac{e^{jk\omega_0 t} - e^{-jk\omega_0 t}}{j2}$$

Let

$$X_k = \frac{a_k}{2} - j\frac{b_k}{2}$$

Since $a_k = a_{-k}$ and $b_k = -b_{-k}$, we have $X_{-k} = X_k^*$, and therefore, we can write

$$x(t) = \sum_{k=-\infty}^{+\infty} X_k e^{jk\omega_0 t} \tag{7.4}$$

where $X_0 = a_0$ and

$$X_k = \frac{1}{T_0} \int_{t_0}^{t_0+T_0} x(t) e^{-jk\omega_0 t} dt \tag{7.5}$$

for any t_0. This representation for $x(t)$ is called the complex FS, and the X_k are called the complex FS coefficients. We say that $x(t)$ and X_k are a FS pair, which is denoted by $x(t) \leftrightarrow X_k$. In general, the FS coefficients X_k are complex. However, if $x(t)$ is an even time function, then X_k is real, which means that $x(t)$ can be represented with only cosine terms, and if $x(t)$ is an odd time function, then X_k is imaginary, which means that $x(t)$ can be represented with only sine terms.

Let us write X_k in polar form to get

$$X_k = \left\| X_k \right\| e^{j\angle X_k}$$

where

$||X_k||$ is the magnitude of X_k

$\angle X_k$ is the angle of X_k

then Equation 7.4 becomes

$$x(t) = X_0 + 2\sum_{k=1}^{\infty} \left\| X_k \right\| \cos(k\omega_0 t + \angle X_k) \tag{7.6}$$

Notice that $||X_k|| = ||X_{-k}||$, an even function of k, and $\angle X_k = -\angle X_{-k}$, an odd function of k. Here we see that $||X_k||$ gives information about the amplitude and $\angle X_k$ gives information about the phase angle of the sinusoidal contribution to $x(t)$ at the frequency $\omega = k\omega_0$. Thus, $||X_k||$ versus k is called the magnitude spectrum of $x(t)$, and $\angle X_k$ versus k is called the phase spectrum of $x(t)$.

As we study cyclical phenomena, we are interested to know the strengths and time displacements of sinusoidal components in $x(t)$ at the frequencies $f = kf_0$ Hz.

7.1.2 Example and Discussion

To be practical, we must truncate the series in Equation 7.4 to use a finite number of terms, giving

$$x(t) \cong \sum_{k=-K}^{+K} X_k e^{jk\omega_0 t} \tag{7.7}$$

To see what happens for different values of K, let us apply Equation 7.7 to the periodic signal shown in Figure 7.1.

The complex FS coefficients for this signal are given by

$$X_k = \frac{1}{3}\int_0^3 x(t)e^{-jk\frac{2\pi}{3}t}\,dt = \frac{1}{3}\int_{0.5}^{2.0} 4e^{-jk\frac{2\pi}{3}t}\,dt = \frac{4}{k\pi}e^{-jk\frac{5\pi}{6}}\,\sin\left(k\frac{\pi}{2}\right), \quad k \neq 0$$

and $X_0 = 2$. Figure 7.2 shows the magnitude spectrum of $x(t)$, which shows the amplitudes of the sinusoidal components of $x(t)$ at the frequencies $f = kf_0$ Hz.

Figure 7.3 shows the application of Equation 7.7 for increasing values of K. Notice the oscillation, called Gibbs oscillation, about points of discontinuity in $x(t)$. Even as K is increased, the oscillation remains with

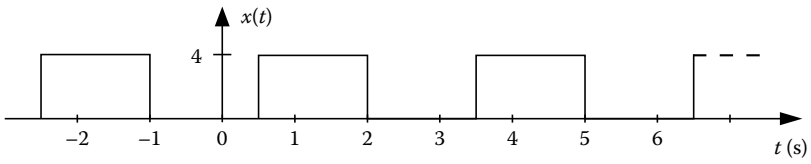


FIGURE 7.1 A periodic signal with discontinuities.

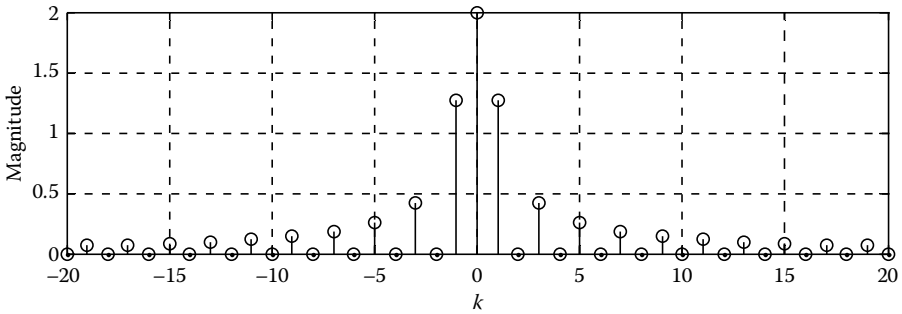


FIGURE 7.2 Magnitude spectrum for $-20 \leq k \leq +20$.

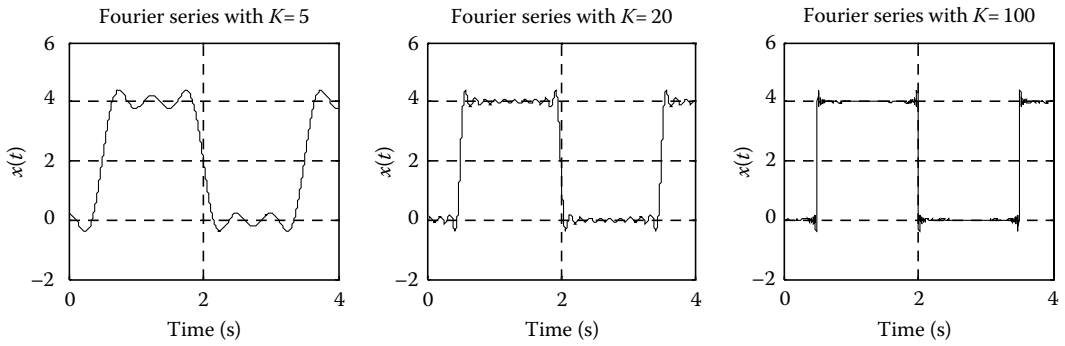


FIGURE 7.3 FS representation with $K = 5$, $K = 20$, and $K = 100$.

an amplitude exceeding $x(t)$ by about 7% of the change in $x(t)$ from one side of the discontinuity to the other side. Also notice that Gibbs oscillation becomes more and more concentrated about the discontinuities as K is increased.

7.1.3 Discrete-Time Signals

We obtain a discrete-time signal by uniformly sampling a continuous-time signal at some sampling rate f_s samples/s (Hz), which is expressed by

$$x(nT) = x(t)|_{t=nT}$$

where n, an integer, is called the discrete-time index and $T = 1/f_s$ is the sample time increment. If T is known and fixed, we may write the discrete-time signal as $x(n)$ instead of $x(nT)$.

Some discrete-time signals are a matter of definition. A few standard signals are

1. Unit step function,
   $u(n - n_0) = \begin{cases} 1, & n - n_0 \geq 0 \\ 0, & n - n_0 < 0 \end{cases}$
2. Unit pulse function (also called the Kronecker delta function),
   $\delta(n - n_0) = \begin{cases} 1, & n - n_0 = 0 \\ 0, & n - n_0 \neq 0 \end{cases}$
3. Exponential function,
   γ^n, for some real or complex constant γ
4. Sinusoidal function,
   $\cos(\theta n + \varphi)$

In terms of a continuous timescale, these discrete-time signals are only defined at the discrete time points given by $t = nT$.

The unit step function is commonly used to start (or stop) a given signal. For example, a sinusoidal pulse can be expressed as $\cos((\pi/4)n)(u(n) - u(n - 8))$, which is one cycle of the sinusoid. The unit pulse function is used to position a value at a time point. For example, $x(n) = -5\delta(n - 2)$ positions the value -5 of x to occur at the time $n = 2$, while for all other n, x has zero value. Through Euler's identity (Equation 7.3), a sinusoidal function can be expressed as the sum of two complex conjugate exponential functions.

We can write any discrete-time signal $x(n)$ as a linear combination of unit pulse functions given by

$$x(n) = \sum_{i=-\infty}^{+\infty} x(i)\delta(n - i) \tag{7.8}$$

7.1.4 Discrete Fourier Transform

For practical signals a function expression for $x(t)$ is usually not available, and therefore we cannot implement Equation 7.5 to find the spectrum X_k of $x(t)$. However, $x(t)$ can be sampled to obtain

$$x(nT) = x(t)|_{t=nT}$$

where T is determined with

$$T = \frac{T_0}{N} \tag{7.9}$$

for some integer N, which makes $x(n)$ a periodic discrete-time signal. We now develop an algorithm to find information about X_k, given $x(n)$.

To process $x(n)$, which can have values in a continuous range, by digital means, we must input each sample $x(n)$ into an analog-to-digital converter to obtain $x_{\mathrm{d}}(n)$, a digital signal, where each sample can have only one of a finite number of possible values. We approximate $x(n)$ with $x_{\mathrm{d}}(n)$, and

$$x(n) = x_{\mathrm{d}}(n) + e_{\mathrm{q}}(n)$$

where $e_{\mathrm{q}}(n)$ is the quantization error. In the following work $x(n)$ will be used for DSP. This may be acceptable if quantization error is negligible.

With the choice for T in Equation 7.9, $x(n)$ is periodic with period N, and with Equation 7.4, $x(n)$ becomes

$$x(n) = \sum_{k=-\infty}^{\infty} X_k \mathrm{e}^{jk\omega_0 nT}$$

Since $\omega_0 = \frac{2\pi}{T_0} = \frac{2\pi}{NT}$, we get for $x(n)$

$$x(n) = \sum_{k=-\infty}^{\infty} X_k \mathrm{e}^{jk\frac{2\pi}{NT}nT} = \sum_{k=-\infty}^{\infty} X_k \mathrm{e}^{j\frac{2\pi}{N}kn}$$

Let us write $k = m + rN$, where $m = 0, 1, \ldots, N-1$ and $r = \ldots, -2, -1, 0, 1, 2, \ldots$ are integers, and then $x(n)$ can be written as

$$x(n) = \sum_{r=-\infty}^{\infty} \sum_{m=0}^{N-1} X_{m+rN} \mathrm{e}^{j\frac{2\pi}{N}(m+rN)n} = \sum_{m=0}^{N-1} \sum_{r=-\infty}^{\infty} X_{m+rN} \mathrm{e}^{j\frac{2\pi}{N}mn}$$

where $\mathrm{e}^{j(2\pi/N)rNn} = 1$. Let $\hat{X}\ (k)$ be determined with

$$\hat{X}(k) = \sum_{r=-\infty}^{\infty} X_{k+rN}$$

and then $x(n)$ becomes

$$x(n) = \sum_{k=0}^{N-1} \hat{X}(k) \mathrm{e}^{j\frac{2\pi}{N}kn} \tag{7.10}$$

Notice that $\hat{X}(k)$ and X_k are related because $\hat{X}(k)$ is the periodic extension of X_k, and therefore $\hat{X}(k)$ is periodic, so that $\hat{X}(k) = \hat{X}(k + iN)$ for any integer i and all k. Knowing X_k, we can find $\hat{X}(k)$, $k = 0, 1, \ldots, N-1$. Since $\hat{X}(k)$ is periodic, we have $\hat{X}(k) = \hat{X}(k-N)$, and since the FS coefficients satisfy $X_k^\star = X_{-k}$, we have $\hat{X}^\star(k) = \hat{X}(N-k)$, and therefore

$$\hat{X}(k) = \hat{X}^\star(N-k), \quad k = 0, 1, \ldots, N-1$$

This means, for example, that $\hat{X}^\star(-1) = \hat{X}(1) = \hat{X}^\star(N-1)$.

It would be very useful if we could obtain $\hat{X}(k)$ directly from $x(n)$. To do this, multiply both sides of Equation 7.10 by a particular exponential function, giving

$$x(n)\mathrm{e}^{-j\frac{2\pi}{N}nm} = \mathrm{e}^{-j\frac{2\pi}{N}nm}\sum_{k=0}^{N-1}\hat{X}(k)\mathrm{e}^{j\frac{2\pi}{N}kn} = \sum_{k=0}^{N-1}\hat{X}(k)\mathrm{e}^{j\frac{2\pi}{N}(k-m)n}$$

where m is any integer in the range $0, 1, \ldots, N-1$. Summing the left-hand and right-hand sides of this equation for $n = 0, 1, \ldots, N-1$ gives

$$\sum_{n=0}^{N-1}x(n)\mathrm{e}^{-j\frac{2\pi}{N}nm} = \sum_{n=0}^{N-1}\sum_{k=0}^{N-1}\hat{X}(k)\mathrm{e}^{j\frac{2\pi}{N}(k-m)n} = \sum_{k=0}^{N-1}\hat{X}(k)\sum_{n=0}^{N-1}\mathrm{e}^{j\frac{2\pi}{N}(k-m)n}$$

Applying the formula for the sum of a geometric series gives

$$\sum_{n=0}^{N-1}\mathrm{e}^{j\frac{2\pi}{N}(k-m)n} = \begin{cases} N, & k = m \\ \dfrac{1-\mathrm{e}^{j\frac{2\pi}{N}(k-m)N}}{1-\mathrm{e}^{j\frac{2\pi}{N}(k-m)}} = 0, & k \neq m \end{cases}$$

and therefore

$$\sum_{n=0}^{N-1}x(n)\mathrm{e}^{-j\frac{2\pi}{N}nm} = N\hat{X}(m)$$

This results in the discrete Fourier transform (DFT) $X(k)$ of $x(n)$, $n = 0, 1, \ldots, N-1$, which is given by

$$X(k) = \sum_{n=0}^{N-1}x(n)\mathrm{e}^{-j\frac{2\pi}{N}nk}, \quad k = 0, 1, \ldots, N-1 \tag{7.11}$$

and therefore

$$X(k) = N\hat{X}(k) = N\sum_{r=-\infty}^{\infty}X_{k+rN} \tag{7.12}$$

Equation 7.10 becomes

$$x(n) = \frac{1}{N}\sum_{k=0}^{N-1}X(k)\mathrm{e}^{j\frac{2\pi}{N}kn} \tag{7.13}$$

which is called the inverse DFT (IDFT). Like $\hat{X}(k)$, $X(k)$ is periodic with period N and $X(k) = X^\star(N-k)$. We write that $X(k) = \mathrm{DFT}\ \{x(n)\}$ and $x(n) = \mathrm{IDFT}\{X(k)\}$.

Equations 7.11 and 7.13 are very similar. If $x(n)$ is real, then conjugating Equation 7.13 gives

$$x(n) = \frac{1}{N}\sum_{k=0}^{N-1} X^*(k)e^{-j\frac{2\pi}{N}kn} = \frac{1}{N}\text{DFT}\{X^*(k)\}$$

Therefore, if we have an algorithm to compute Equation 7.11, the DFT, then that same algorithm, instead of Equation 7.13, can be used to compute the IDFT.

While our interest is to investigate the spectral properties of a continuous-time periodic signal with Equations 7.11 and 7.12, Equation 7.13 is a discrete-time FS representation for a discrete-time periodic signal, and Equation 7.11 is used to find the discrete-time FS coefficients. We say that $x(n)$ and $X(k)$ are a discrete Fourier series (DFS) pair, which is denoted by $x(n) \leftrightarrow X(k)$. Notice that for a discrete-time periodic signal, the FS uses a finite number of discrete-time sinusoids, while for a continuous-time periodic signal, the FS in Equation 7.4 uses a countable infinite number of sinusoids. Furthermore, $X(k)$ is a periodic function in the frequency domain with period N.

7.1.5 Examples and Discussion

Given $x(n)$, the result of sampling a continuous-time periodic signal $x(t)$, we apply Equation 7.11 to obtain $X(k) = \text{DFT}\{x(n)\}$, because, in view of Equation 7.12, over one period of $X(k)$ we expect to obtain a function that potentially behaves like X_k, the spectrum of the continuous-time periodic signal. An important property of $x(t)$ that strongly influences the relationship between X_k and $X(k)$ is whether or not $x(t)$ is bandlimited. If $x(t)$ is bandlimited, then there is some positive integer M such that $X_k = 0$ for $|k| > M$. We say that the bandwidth of $x(t)$ is $\text{BW} = Mf_0$ Hz. Denote the BW by f_c.

To see the effect of selecting an appropriate sampling frequency f_s, let us work with the bandlimited signal given by

$$x(t) = 6\cos\left(10\pi t + \frac{\pi}{4}\right) + 4\sin\left(30\pi t - \frac{\pi}{3}\right) \tag{7.14}$$

Here, $f_0 = 5$ Hz, and $f_c = 15$ Hz, with $M = 3$. Figure 7.4 shows the magnitude spectrum of $x(t)$. The spectral points occur at the frequencies given by $f = k(1/T_0) = kf_0$ Hz, $k = -M, \ldots, 0, 1, 2, M$.

Now, let us sample this $x(t)$ over one period to obtain $N = 10$ time domain samples $x(n)$, $n = 0, 1, \ldots, N-1$. The magnitude of the DFT of $x(n)$ is shown in Figure 7.5. Here we see two periods of the magnitude of $X(k)$ obtained from $||X_k||$ and its first translation $||X_{k-N}||$. Depending on N, the translations of X_k to all integer multiples of N may overlap. In Figure 7.5, N is large enough to prevent the overlap of the translations of X_k, and therefore one period of $X(k)/N$ gives X_k without error. Notice that for $k = 0, 1, \ldots, M$ we have $||X_k||$ for positive frequencies, and for $k = N-1, N-2, \ldots, N-M$ we have $||X_k||$ for negative frequencies.

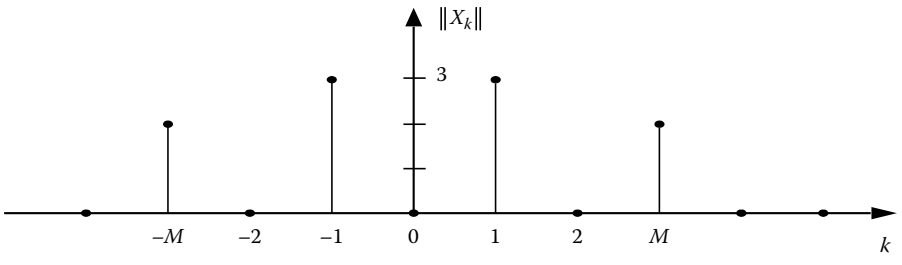


FIGURE 7.4 Magnitude spectrum of a bandlimited signal.

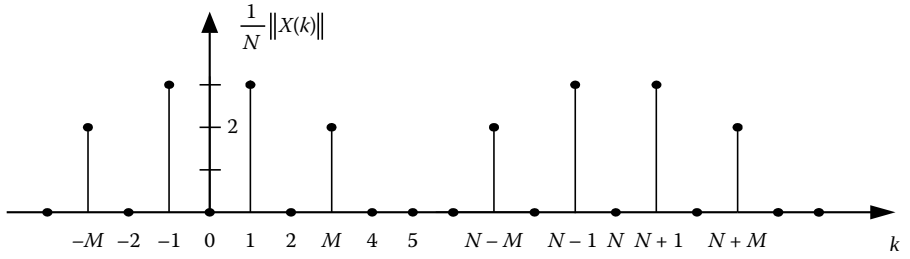


FIGURE 7.5 Two periods of the magnitude of $X(k)$.

In general, if $x(t)$ is bandlimited and $N - M > M$, then the translations of X_k to integer multiples of N cannot overlap. This condition can be written as

$$\frac{N-M}{T_0} > \frac{M}{T_0} \rightarrow \frac{N}{NT} - Mf_0 > Mf_0$$

Therefore, if $x(t)$ is bandlimited to f_c and the sampling frequency f_s satisfies

$$f_s > 2f_c \tag{7.15}$$

such that T_0/T is an integer, which means that f_s must be an integer multiple of f_0, then the DFT of $\{x(n)\}$ gives the spectrum of $x(t)$ without error. This is called the sampling theorem.

In Figure 7.5, the sampling frequency is given by $f_s = \frac{1}{T} = \frac{N}{T_0} = 50$ Hz, and therefore Equation 7.15 is satisfied. Moreover, $x(t)$ can have sinusoidal components with frequencies in the range $0 \leq f < f_s/2$, and the DFT gives X_k without error.

Now, let us try an f_s that does not satisfy Equation 7.15, and such that the frequency points again occur at integer multiples of $f_0 = 5$ Hz. For example, let $f_s = 25$ Hz, giving $N = 5$. Figure 7.6 shows the magnitude of $X(k)$. Like in Figure 7.5, we use the DFT to give the spectrum of $x(t)$ in the frequency range $0 \leq f < f_s/2$, and here $f_s/2 = 12.5$ Hz, which corresponds to $k = 0, 1, 2$. From Figure 7.6, we conclude that $x(t)$ has sinusoidal components at the frequencies 5 and 10 Hz. However, only the spectral point at 5 Hz is correct. The spectral point at 10 Hz, which does not occur in $x(t)$, is called aliasing error, and we cannot even determine that $x(t)$ has a sinusoidal component at 15 Hz. If the original $x(t)$ did consist of sinusoids only at the frequencies 5 and 10 Hz, then Figure 7.6 would be correct.

To obtain correct results with the DFT, we must know the bandwidth and period of a given periodic signal $x(t)$. However, it may not be realistic to assume that we can know this information about a given

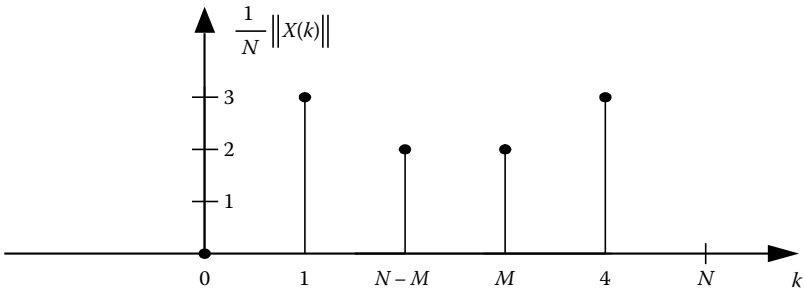


FIGURE 7.6 Magnitude of $X(k)$ for $k = 0, 1, 2, \ldots, N-1$, $N = 5$, and $f_s = 25$ Hz.

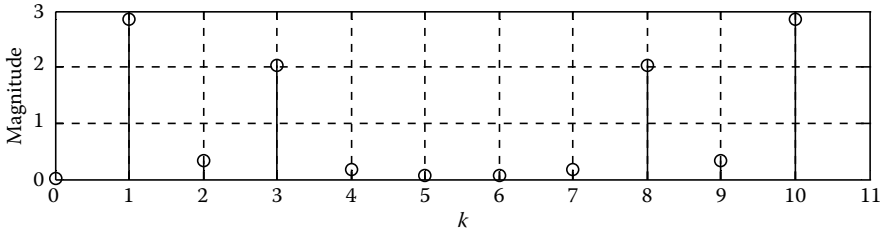


FIGURE 7.7 Magnitude spectrum for $k = 0, 1, \ldots, N-1$, $N = 11$, and $f_s = 53$ Hz.

signal. Furthermore, $x(t)$ may not even be bandlimited (e.g., see Figure 7.2.), in which case translations of X_k to form $X(k)$ will overlap and aliasing error cannot be avoided.

In comparison to the results shown in Figure 7.5, let us try an f_s that does satisfy the sampling theorem of Equation 7.15 for the signal given by Equation 7.14. For example, let $f_s = 53$ Hz. For this sampling frequency, $T_0/T = 10.6$ is not an integer. Therefore, let us set N to $N = 11$, which means that effectively we consider one period $x_p(t)$ to be defined for $0 \leq t < T_0$, where $T_0 = NT = 0.210325$ s. Therefore, the DFT will obtain spectral points at integer multiples of $f_0 = 1/T_0 = 4.7545$ Hz, instead of 5 Hz, as in Figure 7.5. While the signal given by Equation 7.14 is bandlimited, a periodic extension of $x_p(t)$ is a periodic signal with discontinuities at period boundaries, and therefore we have a signal that is not bandlimited. Figure 7.7 shows the magnitude of $X(k)$, and we see that, while it is similar to the magnitude spectrum given in Figure 7.5, this result is an approximation of the spectrum of the signal given by Equation 7.14.

The spectral points in Figure 7.5 at 5 and 15 Hz have spread out about these frequencies in Figure 7.7. This is called leakage error. Leakage error occurs when we use a sample time range other than an integer multiple of the period of the given periodic signal.

If it is practically possible, it is useful to sample the continuous-time signal over a longer time range. For example, with the sampling frequency used to produce the result shown in Figure 7.7, let us obtain $N = 50$ samples of $x(t)$. In this case, $x_p(t)$ becomes the segment of $x(t)$ for $0 \leq t < T_0$, where $T_0 = NT = 0.9434$ s. It is likely that we will still have discontinuities at the period boundaries of the periodic extension of $x_p(t)$, which means that the resulting periodic signal is not bandlimited, and aliasing error cannot be avoided. Since T_0 is not an integer multiple of the period of the given $x(t)$, we will incur leakage error. However, we now have a higher frequency resolution determined by $f_0 = 1/T_0 = 1.06$ Hz.

Figure 7.8 shows the magnitude of the DFT of $x(n)$, $n = 0, 1, \ldots, 49$. Notice that the spectral point at $k = 0$ is not zero, which means that the average value of the periodic signal is not zero, even though the

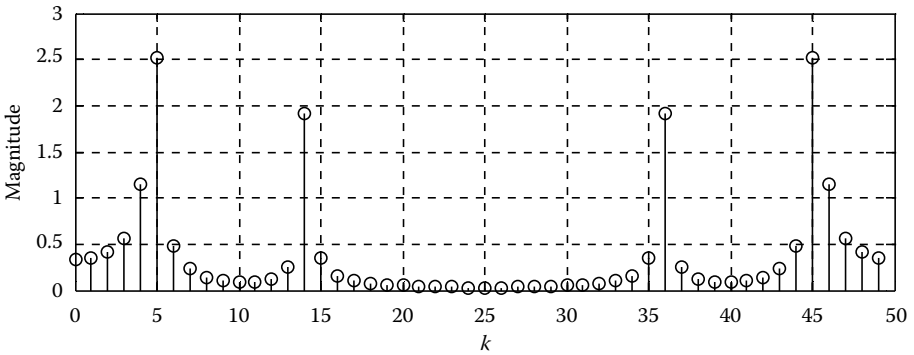


FIGURE 7.8 Magnitude spectrum for $k = 0, 1, \ldots, N-1$, $N = 50$, and $f_s = 53$ Hz.

given $x(t)$ has a zero average value. Also, there is leakage error, as expected, and there is significant aliasing error at higher frequencies, i.e., as $k \rightarrow 25$. Nevertheless, like in Figure 7.7, this magnitude spectrum is an approximation of the magnitude spectrum shown in Figure 7.5.

We can reduce the aliasing error shown in Figure 7.8 by reducing the discontinuities that occur at the period boundaries of the periodic extension of $x_p(t)$. Let us view $x_p(t)$ to be given by

$$x_p(t) = x(t)w(t) \tag{7.16}$$

where $w(t)$, which is called a window function, is given by $w(t) = w_R(t)$, the rectangular window, and

$$w_R(t) = \begin{cases} 1, & 0 \leq t < T_0 \\ 0, & \text{otherwise} \end{cases}$$

We can write $w_R(t) = u(t) - u(t - T_0)$, where $u(t)$ is the unit step function.

To eliminate the discontinuities at period boundaries, let the window function be given by $w(t) = w_H(t)$, where

$$w_H(t) = \frac{1}{2}\left(1 - \cos\left(\frac{2\pi t}{T_0}\right)\right) w_R(t)$$

which is called the Hann window. Notice that $w_H(t = 0) = w_H(t = T_0) = 0$. Using the Hann window in Equation 7.16 gives the magnitude spectrum shown in Figure 7.9. Notice the reduced magnitudes about the frequencies of sinusoids in the given $x(t)$. There are numerous other widow functions that could also have been used.

7.1.6 DFT Applications

The DFT has many properties that account for its wide application. Here, we examine how the DFT can be involved to compute the linear convolution of two discrete-time signals. Let $x_1(n)$ and $x_2(n)$ be two periodic discrete-time signals, each having period N. The circular convolution of $x_1(n)$ and $x_2(n)$, which is denoted by $x_1(n) \otimes x_2(n)$, is another periodic signal $x_3(n)$ with period N given by

$$x_3(n) = x_1(n) \otimes x_2(n) = \sum_{i=0}^{N-1} x_1(i)x_2(n-i) = \sum_{i=0}^{N-1} x_1(n-i)x_2(i) \tag{7.17}$$

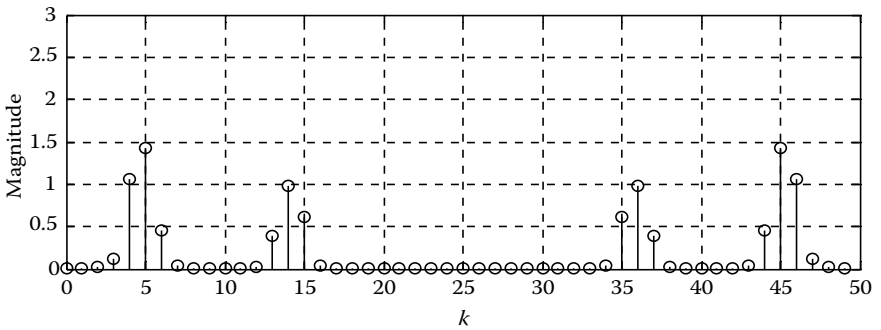


FIGURE 7.9 Magnitude spectrum for $N = 50$, $f_s = 53$ Hz, and $w(t) = w_H(t)$.

It is not difficult to show that $X_3(k) = \text{DFT}\{x_3(n)\}$ is given by

$$X_3(k) = X_1(k)X_2(k) \tag{7.18}$$

which is called the circular convolution theorem. In other words, circular convolution in the time domain can be transformed to multiplication in the frequency domain. Therefore, instead of using Equation 7.17 to find the circular convolution of $x_1(n)$ and $x_2(n)$, we find $x_3(n)$ with the IDFT of $X_3(k)$ found with Equation 7.18. This is the indirect convolution method.

Now consider the linear convolution of $y_1(n)$ and $y_2(n)$, which are not periodic, to obtain $y_3(n)$ given by

$$y_3(n) = \sum_{i=-\infty}^{+\infty} y_1(n-i)y_2(i) = \sum_{i=-\infty}^{+\infty} y_1(i)y_2(n-i) = y_1(n)^* y_2(n) \tag{7.19}$$

Let us assume that $y_1(n)$ and $y_2(n)$ have finite durations, and $y_1(n)$ is given for $n = 0, \ldots, N_1 - 1$ and $y_2(n)$ is given for $n = 0, \ldots, N_2 - 1$. The duration N_3 of $y_3(n)$ will be $N_3 = N_1 + N_2 - 1$.

To employ the DFT, let N be any integer such that $N > N_3$, and define periodic signals $x_1(n)$ and $x_2(n)$ with periods given by

$$x_1(n) = \begin{cases} y_1(n), & n = 0, \ldots, N_1 - 1 \\ 0, & n = N_1, \ldots, N - 1 \end{cases} \quad \text{and} \quad x_2(n) = \begin{cases} y_2(n), & n = 0, \ldots, N_2 - 1 \\ 0, & n = N_2, \ldots, N - 1 \end{cases}$$

We now find the circular convolution $x_3(n)$ of $x_1(n)$ and $x_2(n)$ with the indirect convolution method, and then over one period of $x_3(n)$ we get $y_3(n) = x_3(n)$ for $n = 0, \ldots, N_3 - 1$. At this point, we may expect that the indirect convolution method will require more computation than the direct convolution method using Equation 7.17. However, we will develop a computationally efficient algorithm, called the fast Fourier transform (FFT) to compute the DFT that will make the indirect convolution method more computationally efficient than the direct convolution method.

An operation similar to linear convolution is the linear correlation operation given by

$$r_{12}(l) = \sum_{n=-\infty}^{+\infty} y_1(n)y_2(n-l) \tag{7.20}$$

where l, $l = 0, 1, \ldots$, is called the lag index. However, the correlation $r_{12}(l)$, which is useful to investigate the similarity of two signals, can be obtained with

$$r_{12}(l) = y_1(l)^* y_2(-l) \tag{7.21}$$

Again, we assume that $y_1(n)$ and $y_2(n)$ have finite durations of N_1 and N_2, respectively, and the duration of $r_{12}(l)$ will be $N_3 = N_1 + N_2 - 1$.

To employ the DFT, we define $x_1(n)$ and $x_2(n)$ as before, and by indirect convolution compute

$$x_3(l) = x_1(l) \otimes x_2(-l)$$

to get $r_{12}(l) = x_3(l)$ for $l = 0, \ldots, N_3 - 1$. The DFT of $x_2(-n)$ is given by $X_2^*(k)$, where $X_2(k)$ is the DFT of $x_2(n)$.

7.1.7 Fast Fourier Transform

The FFT is a collection of algorithms that can compute the DFT, $X(k)$, $k = 0, \ldots, N-1$, of a discrete-time signal, $x(n)$, $n = 0, \ldots, N-1$, much more efficiently than with Equation 7.11. The efficiency of the FFT occurs when N can be expressed as a product of many small integers, $N = N_1\, N_2 \ldots N_I$. An ideal case occurs when N is a power of 2, $N = 2^L$, for some integer L, and then the FFT algorithm is called the radix-2 FFT algorithm. Other possibilities are, for example, $N = 3^K$, resulting in a radix-3 FFT, or N is a product of many different small integers, resulting in a mixed-radix FFT. Here, we will outline the development of the radix-2 FFT, and even more specifically, the decimation in time radix-2 FFT.

Before we start the FFT development, let us assess the computational requirements of Equation 7.11, the DFT operation. It is generally assumed that sinusoidal values are found by table lookup, which requires time that is negligible compared to multiplication time. If the DFT input is a set of N complex numbers, then the DFT will require a total of N^2 complex multiplications to obtain the set of N outputs, and we will use this for assessment of DFT computational requirements. We will see how the radix-2 FFT works to require only $NL/2$ complex multiplications.

We start by assuming that the number of points is $N = 2^L$, for some integer L. If for a given discrete-time signal $x(n)$ the number of points N is not a power of two, then we can augment the given $x(n)$ with a sufficient number of zero points for a total number of points that is a power of 2.

Since N is even, we can split the DFT summation in Equation 7.11 into two parts having an equal number of summation terms. Let the first part come from the even indexed terms in Equation 7.11, and let the second part come from the odd indexed terms, and we get

$$X(k) = \sum_{n=0}^{N-1} x(n) \mathrm{e}^{-j\frac{2\pi}{N}nk} = \sum_{n,\text{even}} x(n) \mathrm{e}^{-j\frac{2\pi}{N}nk} + \sum_{n,\text{odd}} x(n) \mathrm{e}^{-j\frac{2\pi}{N}nk}$$

Even n can be written as $n = 2i$, and odd n can be written as $n = 2i + 1$, for $i = 0, 1, \ldots, N_1 - 1$, where $N_1 = N/2$, which is an even integer. Now we have

$$X(k) = \sum_{i=0}^{N_1-1} x(2i) \mathrm{e}^{-j\frac{2\pi}{N}(2i)k} + \sum_{i=0}^{N_1-1} x(2i+1) \mathrm{e}^{-j\frac{2\pi}{N}(2i+1)k}$$

Let $x_0(n)$ and $x_1(n)$ be the N_1 point number sequences defined by

$$\begin{aligned} x_0(n) &= x(2n) \\ x_1(n) &= x(2n+1), \quad n = 0, 1, \ldots, N_1 - 1 \end{aligned}$$

and therefore we get

$$X(k) = \sum_{n=0}^{N_1-1} x_0(n) \mathrm{e}^{-j\frac{2\pi}{N_1}nk} + \mathrm{e}^{-j\frac{2\pi}{N}k} \sum_{n=0}^{N_1-1} x_1(n) \mathrm{e}^{-j\frac{2\pi}{N_1}nk}$$

Now let

$$X_0(k) = \sum_{n=0}^{N_1-1} x_0(n) \mathrm{e}^{-j\frac{2\pi}{N_1}nk} \tag{7.22}$$

$$X_1(k) = \sum_{n=0}^{N_1-1} x_1(n) \mathrm{e}^{-j\frac{2\pi}{N_1}nk}, \quad k = 0, 1, \ldots, N_1 - 1 \tag{7.23}$$

and therefore we get

$$X(k) = X_0(k) + e^{-j\frac{2\pi}{N}k}X_1(k), \quad k = 0, 1, \ldots, N-1 \tag{7.24}$$

An N-point DFT, $X(k)$, is periodic with period N. We see that Equations 7.22 and 7.23 are each N_1-point DFTs, and therefore $X_0(k)$ and $X_1(k)$ are periodic with period N_1. An N_1-point DFT requires N_1^2 complex multiplications. To find $X(k)$, let us use Equation 7.24 only for $k = 0, 1, \ldots, N_1$. However, for $k = N_1, N_1 + 1, \ldots, N-1$, let $m = k - N_1$, so that m varies from $m = 0$ to $m = N_1 - 1$, and Equation 7.24 becomes

$$X(N_1 + m) = X_0(N_1 + m) + e^{-j\frac{2\pi}{N}(N_1+m)}X_1(N_1 + m)$$

Or, since $X_0(k)$ and $X_1(k)$ are periodic with period N_1, we have

$$X(N_1 + k) = X_0(k) - e^{-j\frac{2\pi}{N}k}X_1(k), \quad k = 0, 1, \ldots, N_1 - 1 \tag{7.25}$$

The only difference between Equations 7.24 and 7.25 is addition of two terms in Equation 7.24 and subtraction of the same two terms in Equation 7.25.

Figure 7.10 combines Equations 7.24 and 7.25 for $k = 0, 1, \ldots, N_1 - 1$. This structure is called a butterfly operation. To find $X(k)$ for $k = 0, 1, \ldots, N-1$ given $X_0(k)$ and $X_1(k)$ for $k = 0, 1, \ldots, N_1 - 1$ requires $N/2$ complex multiplications, one complex multiplication for each butterfly operation. The $N/2$ butterfly operations are the first stage of reducing the computational requirements of the DFT operation.

The sequence of steps toward the first stage started with splitting the summation of Equation 7.11 by decimating $x(n)$ and finished with Equations 7.24 and 7.25, which is illustrated by Figure 7.10. We continue to obtain the second stage by decimating $x_0(n)$ into $x_{00}(n)$ and $x_{10}(n)$ from the even and odd, respectively, indexed points in $x_0(n)$ and by decimating $x_1(n)$ into $x_{01}(n)$ and $x_{11}(n)$ from the even and odd, respectively, indexed points in $x_1(n)$ to reduce the computational requirements for $X_0(k)$ and $X_1(k)$. This second stage requires $N_1/2 + N_1/2 = N/2$ complex multiplications. We continue through L stages, where in the last stage, each of the $N/2$ DFT operations is a 2-point DFT consisting of a single butterfly operation. Therefore, the last stage also requires $N/2$ complex multiplications. The overall algorithm is called the decimation in time radix-2 FFT, and it requires $NL/2$ complex multiplications to compute the DFT, compared to a direct DFT computation, which requires N^2 complex multiplications.

Suppose a signal $x(t)$ with BW = 3.6 kHz is sampled at the rate $f_s = 8$ kHz to obtain $N = 2^{13}$ points (about a 1 s time interval). A direct DFT requires approximately 67,000,000 complex multiplications, while an FFT requires approximately 53,000 complex multiplications, a reduction in time by a factor of 1264. The efficiency of the FFT improves as N is increased. With the FFT, indirect convolution can be demonstrated to be much more efficient than direct convolution.

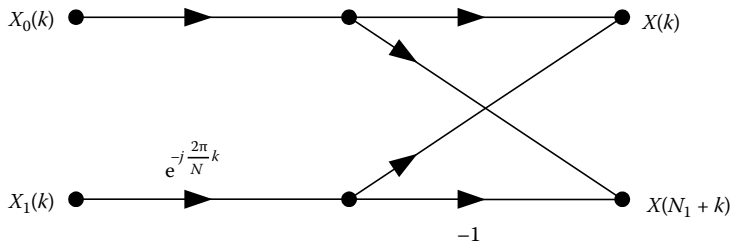


FIGURE 7.10 Butterfly operations of the first stage.

7.1.8 Continuous-Time Fourier Transform

By sampling a periodic continuous-time signal we developed the DFT operation and found a relationship between the result of a DFT operation $X(k)$ and the spectrum X_k of the continuous-time periodic signal. We also came to understand how $X(k)$ can be different from X_k, the FS coefficients. However, real-world continuous-time signals are not necessarily periodic. Here, we extend the concept of a FS approximation of a periodic continuous-time signal to an aperiodic continuous-time signal.

Consider the arbitrary aperiodic signal shown in Figure 7.11, where T_0 is large enough to contain the signal within the time range $-T_0/2 < t < T_0/2$. Over this time range the given signal can be represented with the FS of Equation 7.4, which is repeated here for convenience

$$x(t) = \sum_{k=-\infty}^{+\infty} X_k e^{jk\omega_0 t} \tag{7.26}$$

where

$$X_k = \frac{1}{T_0} \int_{-T_0/2}^{+T_0/2} x(t) e^{-jk\omega_0 t} dt \tag{7.27}$$

If in Figure 7.11 we let T_0 go to infinity, then the FS in Equation 7.26 represents the given signal for all t. Let us see what happens as T_0 is increased. If the integral in Equation 7.27 remains finite, then the X_k will go to zero. Since the spectral points occur at the frequencies given by $\omega = k\omega_0$, where $\omega_0 = 2\pi/T_0$, the spectrum tends toward a continuous function of frequency. Therefore, in view of Equation 7.27, let us work with

$$X(j\omega) = \lim_{T_0 \to \infty} X_k T_0 \tag{7.28}$$

which gives

$$X(j\omega) = \int_{-\infty}^{+\infty} x(t) e^{-j\omega t} dt \tag{7.29}$$

where $k\omega_0$ has become the continuous frequency ω. And then, Equation 7.26 becomes

$$x(t) = \frac{1}{2\pi} \lim_{T_0 \to \infty} \sum_{k=-\infty}^{+\infty} X_k T_0 e^{jk\omega_0 t} \frac{2\pi}{T_0}$$

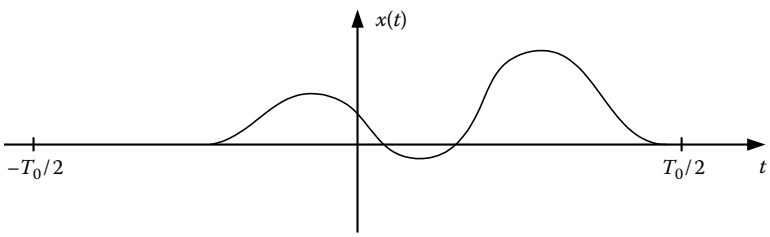


FIGURE 7.11 An arbitrary aperiodic signal.

TABLE 7.1 FT Properties

Linearity	$c_1x_1(t)+c_2x_2(t) \leftrightarrow c_1X_1(j\omega)+c_2X_2(j\omega)$
Time shift	$x(t-t_0) \leftrightarrow X(j\omega)e^{-j\omega t_0}$
Modulation	$x(t)e^{j\omega_0 t} \leftrightarrow X(j(\omega-\omega_0))$
Linear convolution	$x_1(t)^*x_2(t) = \int_{-\infty}^{+\infty} x_1(t-\tau)x_2(\tau)d\tau \leftrightarrow X_1(j\omega)X_2(j\omega)$
Derivative	$\frac{d}{dt}x(t) \leftrightarrow j\omega X(j\omega)$
Multiplication	$x_1(t)x_2(t) \leftrightarrow \frac{1}{2\pi}X_1(j\omega)^*X_2(j\omega)$
Energy (Parseval's theorem)	$\int_{-\infty}^{+\infty} \lvert x(t)\rvert^2 dt = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \lVert X(j\omega)\rVert^2 d\omega$

which gives

$$x(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} X(j\omega)e^{j\omega t}d\omega \tag{7.30}$$

We should view Equation 7.30 similar to Equation 7.26, which is that Equation 7.30 expresses an aperiodic continuous-time signal in terms of a linear combination of an uncountable infinite number of sinusoids. Equation 7.29 gives the Fourier transform (FT) $X(j\omega)$, which is a complex function of the continuous real frequency variable ω, of the continuous-time function $x(t)$, and it describes how the strengths of sinusoidal components of $x(t)$ are distributed over frequency. In view of Equation 7.28, $X(j\omega)$ is the spectral density of $x(t)$. Equation 7.30 gives the inverse FT of $X(j\omega)$, and we say that $x(t)$ and $X(j\omega)$ are a FT pair, denoted by $x(t) \leftrightarrow X(j\omega)$.

The FT has many properties that are widely applied in continuous-time signal and system analysis and design. Assume that we have the FT pair $x(t) \leftrightarrow X(j\omega)$. Table 7.1 gives a few important properties.

7.1.9 Examples and Discussion

The FT of a signal may not exist. For example, for $x(t) = u(t)$, the unit step function, we must utilize an impulse function to express its FT. In view of Equation 7.28, the FT of a sinusoidal signal also requires using impulse functions. A sufficient condition for the existence of the FT is that

$$\lVert X(j\omega)\rVert \le \int_{-\infty}^{+\infty} \left\lVert x(t)e^{-j\omega t}\right\rVert dt = \int_{-\infty}^{+\infty} \lVert x(t)\rVert dt < \infty$$

which means that the signal $x(t)$ must be absolutely integrable.

Consider the signal $x(t)$ given by $x(t) = \sin(\omega_0 t)w_R(t)$, where $w_R(t)$ is a rectangular window extending from $-\tau/2$ to $+\tau/2$, and $x(t)$ is a sinusoidal pulse. To understand the reasons for the features of $X(j\omega)$, the FT of $x(t)$, let us first consider $w_R(t)$. The FT of $w_R(t)$ is given by

$$W_R(j\omega) = \int_{-\tau/2}^{+\tau/2} e^{-j\omega t}dt = \frac{1}{-j\omega}(e^{-j\omega\tau/2} - e^{+j\omega\tau/2}) = \tau\frac{\sin(\omega\tau/2)}{\omega\tau/2}$$

which is a real function of ω, since $w_R(t)$ is an even function, and its magnitude is shown in Figure 7.12. The first zeros of $W_R(j\omega)$ occur at $\omega = \pm 2\pi/\tau$, and within this frequency range we have a feature called the main lobe of $W_R(j\omega)$. Outside this frequency range are the side lobes of $W_R(j\omega)$. We see that as

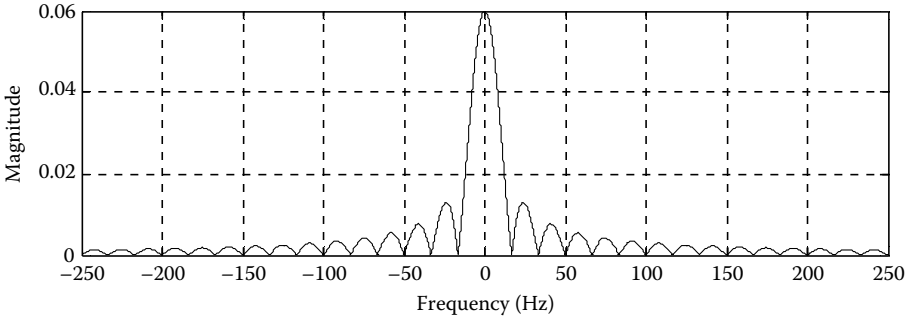


FIGURE 7.12 Spectral density magnitude, $\tau = 0.06$ s.

τ is increased, which makes $w_R(t)$ occur over a longer time range, the main lobe narrows and more energy of $w_R(t)$ is concentrated about $\omega = 0$ and within a smaller frequency range. Conversely, as τ is decreased, which makes $w_R(t)$ occur over a shorter time range, the main lobe broadens and the energy of $w_R(t)$ is spread out over a wider frequency range.

In $x(t)$, let $\omega_0 = 100\pi$ rad/s. The spectral density $X(j\omega)$ of $x(t)$ can be found with the modulation property, which translates $W_R(j\omega)$ up to $\omega = +\omega_0$ and down to $\omega = -\omega_0$. This is shown in Figure 7.13, where we see the main lobes positioned about $\omega = \pm\omega_0$.

In view of the properties given in Table 7.1, the FT is very well suited to investigate the spectral nature of a signal. It is also useful to study the frequency selective behavior of a linear and time invariant (LTI) continuous-time system. Let $h(t)$ denote the impulse response of an LTI system. If the system is initially at rest, then the response $y(t)$ to an input $x(t)$ is given by

$$y(t) = h(t)^* x(t) = \int_{-\infty}^{+\infty} h(t-\tau)x(\tau)\mathrm{d}\tau \tag{7.31}$$

and if the FT $H(j\omega)$ of the impulse response $h(t)$ exists, then by the convolution property we have

$$Y(j\omega) = H(j\omega)X(j\omega)$$

where $H(j\omega)$ is called the system transfer function. Therefore, with $H(j\omega)$ we can find how the system modifies the spectral density of the input to obtain the spectral density of the output. Furthermore, if the input is a sinusoid given by

$$x(t) = A\cos(\omega t + \phi)u(t)$$

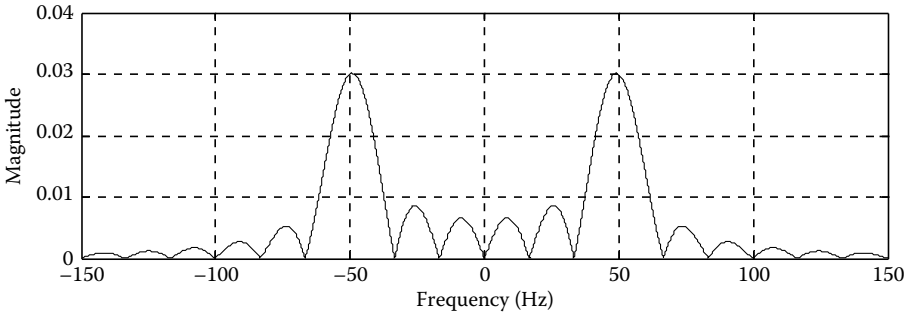


FIGURE 7.13 Spectral density of a sinusoidal pulse.

then the steady-state response of the system will be

$$y(t) = A\|H(j\omega)\| \cos(\omega t + \phi + \angle H(j\omega))$$

which shows the frequency-selective behavior of the system. We plot $||H(j\omega)||$ and $\angle H(j\omega)$ versus ω to see the frequency response of the system.

7.1.10 Discrete-Time Fourier Transform

As for continuous-time periodic signals, we want to study continuous-time aperiodic signals and continuous-time LTI systems by digital means. We start by sampling $x(t)$, the inverse FT of $X(j\omega)$, at the sampling rate $f_s = 1/T$ Hz, and Equation 7.30 becomes

$$x(nT) = \int_{-\infty}^{+\infty} X(j2\pi f) e^{j2\pi f nT} \mathrm{d}f$$

Breaking the frequency range into intervals of length f_s results in

$$x(nT) = \sum_{r=-\infty}^{+\infty} \int_{(2r-1)\frac{f_s}{2}}^{(2r+1)\frac{f_s}{2}} X(j2\pi f) e^{j2\pi f nT} \mathrm{d}f$$

Let us change the integration variable from f to $\lambda = f - rf_s$, and then we have

$$x(nT) = \sum_{r=-\infty}^{+\infty} \int_{-\frac{f_s}{2}}^{+\frac{f_s}{2}} X(j2\pi(\lambda + rf_s)) e^{j2\pi(\lambda + rf_s)nT} \mathrm{d}\lambda$$

Interchanging summation and integration and simplifying the exponential term results in

$$x(nT) = \int_{-\frac{f_s}{2}}^{+\frac{f_s}{2}} \sum_{r=-\infty}^{+\infty} X(j2\pi(\lambda + rf_s)) e^{j2\pi\lambda nT} \mathrm{d}\lambda$$

Reverting to f for the frequency variable gives

$$x(nT) = \int_{-\frac{f_s}{2}}^{+\frac{f_s}{2}} \sum_{r=-\infty}^{+\infty} X(j2\pi(f - rf_s)) e^{j2\pi f nT} \mathrm{d}f \tag{7.32}$$

Let $\hat{X}(j2\pi f)$ be determined with

$$\hat{X}(j2\pi f) = \sum_{r=-\infty}^{+\infty} X(j2\pi(f - rf_s)) \tag{7.33}$$

which is the periodic extension of $X(j2\pi f)$, and therefore $\hat{X}(j2\pi f)$ is a periodic function of frequency with period f_s. And, Equation 7.32 becomes

$$x(nT) = \int_{-\frac{f_s}{2}}^{+\frac{f_s}{2}} \hat{X}(j2\pi f) e^{j2\pi f nT} df \tag{7.34}$$

If we know $X(j2\pi f)$, the FT of the real-time function $x(t)$, then we can find $\hat{X}(j2\pi f)$ with Equation 7.33. It would be useful if we could obtain $\hat{X}(j2\pi f)$ directly from the samples $x(nT)$ of $x(t)$. Since $\hat{X}(j2\pi f)$ is a periodic function, having an even magnitude and odd phase, we can write it as a FS, which is

$$\hat{X}(j2\pi f) = \sum_{n=-\infty}^{+\infty} \hat{x}_n e^{-jn\frac{2\pi}{f_s}f} \tag{7.35}$$

where the $\hat{x}_n$ are the FS coefficients given by

$$\hat{x}_n = \frac{1}{f_s} \int_{-\frac{f_s}{2}}^{+\frac{f_s}{2}} \hat{X}(j2\pi f) e^{jn\frac{2\pi}{f_s}f} df \tag{7.36}$$

In view of Equation 7.34, we have $\hat{x}_n = Tx(nT)$. Based on Equation 7.35, let

$$X(e^{j\omega T}) = \sum_{n=-\infty}^{+\infty} x(nT) e^{-j\omega nT} \tag{7.37}$$

which is called the discrete-time Fourier transform (DTFT) of $x(nT)$. Then, $\hat{X}(j\omega) = TX(e^{j\omega T})$, and Equation 7.34 becomes

$$x(nT) = \frac{1}{\omega_s} \int_{-\frac{\omega_s}{2}}^{+\frac{\omega_s}{2}} X(e^{j\omega T}) e^{j\omega nT} d\omega \tag{7.38}$$

which is called the IDTFT. A sufficient condition for the existence of the DTFT is

$$\left\| X(e^{j\omega T}) \right\| = \left\| \sum_{n=-\infty}^{+\infty} x(nT) e^{-j\omega nT} \right\| \leq \sum_{n=-\infty}^{+\infty} \|x(nT)\| < \infty$$

which means that $x(nT)$ must be absolutely summable. We say that $x(nT)$ and $X(e^{j\omega T})$ are a DTFT pair, denoted by $x(nT) \leftrightarrow X(e^{j\omega T})$.

With Equation 7.33, we have

$$X(e^{j\omega T}) = \frac{1}{T} \sum_{r=-\infty}^{+\infty} X(j(\omega - r\omega_s)) \tag{7.39}$$

With Equation 7.39 we can find information about the spectral density of $x(t)$ from the DTFT of $x(nT)$.

7.1.11 DTFT Model

The DTFT is often defined based on the block diagram shown in Figure 7.14, where the impulse sampler is an impulse train $\delta_T(t)$ given by

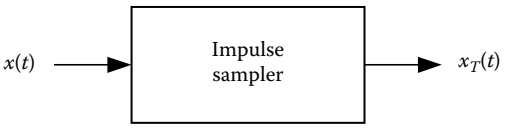


FIGURE 7.14 Model of the sampling process.

$$\delta_T(t) = \sum_{n=-\infty}^{+\infty} \delta(t - nT)$$

and $\delta(t)$ is the Dirac impulse function. The impulse sampler output is

$$x_T(t) = x(t)\delta_T(t) = \sum_{n=-\infty}^{+\infty} x(nT)\delta(t - nT) \tag{7.40}$$

and the FT of $x_T(t)$ gives Equation 7.37, which is the DTFT of $x(nT)$.

7.1.12 Sampling Theorem

Suppose a signal $x(t)$ is bandlimited. Then, there is some ω_c such that $X(j\omega) = 0$ for $|\omega| > \omega_c$. Let us use the spectral density shown in Figure 7.15 to illustrate what happens when we apply the DTFT.

If we sample $x(t)$, then according to Equation 7.39 the DTFT of $x(nT)$ is the periodic extension of $X(j\omega)$, as shown in Figure 7.16. Here, the sampling frequency is high enough to prevent translations of $X(j\omega)$ from overlapping, which means that the spectral density $X(j\omega)$ is contained entirely within the frequency range $(-\omega_s/2, +\omega_s/2)$ of the DTFT of $x(nT)$. Therefore, if

$$\omega_s > 2\omega_c \tag{7.41}$$

then $X(j\omega)$ can be recovered from one period of $X(e^{j\omega T})$ without error.

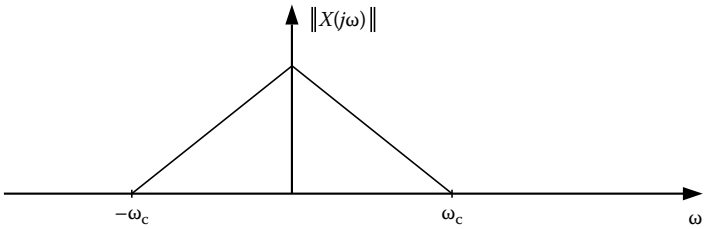


FIGURE 7.15 Spectral density magnitude of a bandlimited signal.

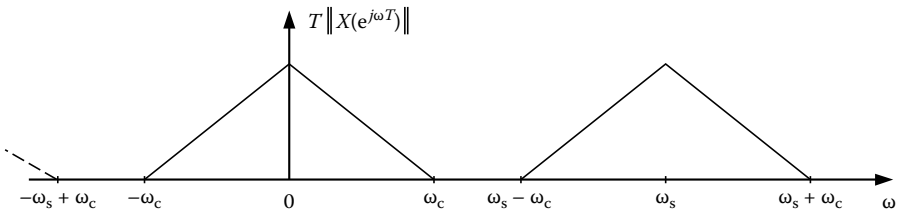


FIGURE 7.16 Magnitude of the DTFT.

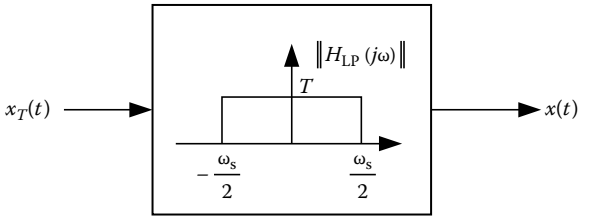


FIGURE 7.17 Ideal low-pass filter magnitude frequency response.

Since, for example, Figure 7.16 also shows the magnitude of the FT of $x_T(t)$, according to Equation 7.40, we can obtain the signal $x(t)$, which has the spectral density shown in Figure 7.15, if we process $x_T(t)$ with an analog low-pass filter as shown in Figure 7.17.

We will find $x(t)$ by convolving the input with the impulse response of the ideal low-pass filter. The impulse response $h_{LP}(t)$ is given by

$$h_{LP}(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} H_{LP}(j\omega)e^{j\omega t}d\omega = \frac{1}{2\pi}\int_{-\omega_s/2}^{+\omega_s/2} Te^{j\omega t}d\omega = \frac{\sin\left(\frac{\omega_s}{2}t\right)}{\frac{\omega_s}{2}t}$$

and then we get

$$x(t) = h_{LP}(t)^* x_T(t) = \sum_{n=-\infty}^{+\infty} x(nT)\frac{\sin\left(\frac{\omega_s}{2}(t-nT)\right)}{\frac{\omega_s}{2}(t-nT)} \tag{7.42}$$

which is called the data reconstruction formula, to obtain $x(t)$, for all t, from the samples $x(nT)$. Together, Equations 7.41 and 7.42 are called the time sampling theorem.

If the sampling frequency ω_s does not satisfy Equation 7.41, then the DTFT produces results like those shown in Figure 7.18. Here we see translations of $X(j\omega)$ overlap. In the frequency range $\omega_s - \omega_c$ to $\omega_s/2$, where $\omega_c > \omega_s/2$, the DTFT of $x(nT)$ is not the same as $X(j\omega)$. This difference is called aliasing error. If we reproduce an analog signal with Equation 7.42, then this analog signal will have energy in the frequency range $\omega_s - \omega_c$ to $\omega_s/2$ that is not the same as that of the given signal $x(t)$. In this situation, if the sampling frequency ω_s cannot be increased to avoid aliasing error, then the given signal should be processed by a low-pass filter, called an antialiasing filter, with bandwidth equal to $\omega_s/2$, which becomes the effective bandwidth of the signal that will be sampled. If the resulting samples are used to reconstruct an analog signal, then the resulting analog signal will not have the very high frequency spectral content of the given signal $x(t)$ in the frequency range $\omega_s/2$ to ω_c.

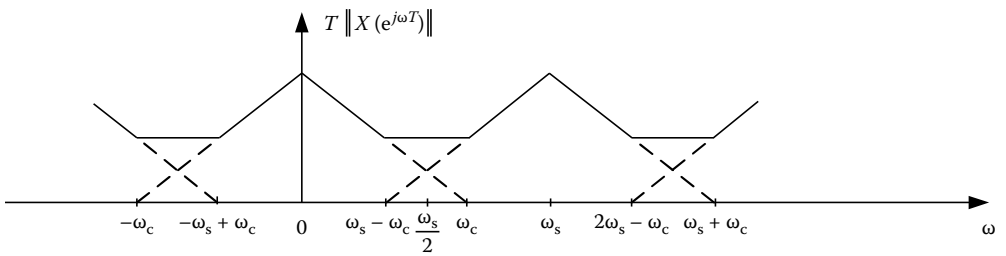


FIGURE 7.18 Magnitude of the DTFT, $\omega_s < 2\omega_c$.

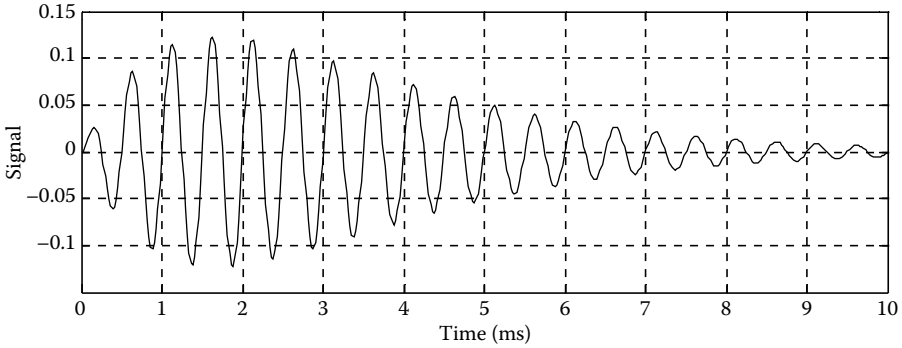


FIGURE 7.19 An exponentially weighted sinusoidal pulse.

If an analog signal $x(t)$ is not bandlimited, then aliasing error cannot be avoided. An antialiasing filter could be used at the cost of eliminating the high frequency sinusoidal components of the signal. Let us use the signal $x(t) = (e^{-at} - e^{-bt})\sin(2\pi f_0 t)u(t)$, where $a = 500$, $b = 700$, and $f_0 = 2000$ Hz, to illustrate what happens. This signal is shown in Figure 7.19. We see that $x(t)$ is a tone that gains strength and then loses strength. The FT of $e^{-at}u(t)$ is $1/(j\omega + a)$, and then the modulation property can be used to find the FT of $x(t)$.

The DTFT of $e^{-anT}u(n)e^{-j\omega_0 nT} = (e^{-aT}e^{-j\omega_0 T})^n u(n)$ is given by

$$\sum_{n=0}^{+\infty} e^{-(aT+j\omega_0 T)n}e^{-j\omega nT} = \lim_{N\to\infty}\sum_{n=0}^{N-1}(e^{-(a+j\omega_0+j\omega)T})^n = \lim_{N\to\infty}\frac{1-(e^{-(a+j\omega_0+j\omega)T})^N}{1-e^{-(a+j\omega_0+j\omega)T}}$$
$$= \frac{1}{1-e^{-aT}e^{-j(\omega+\omega_0)T}} \tag{7.43}$$

The magnitudes of $X(j\omega)$ and $X(e^{j\omega T})$ are shown in Figure 7.20, where they are shown normalized to show them on the same scale. Within the frequency range 0 to $f_s/2$ we see aliasing error that increases as the frequency increases toward $f_s/2$. If we increase the sampling frequency, then the aliasing error will decrease.

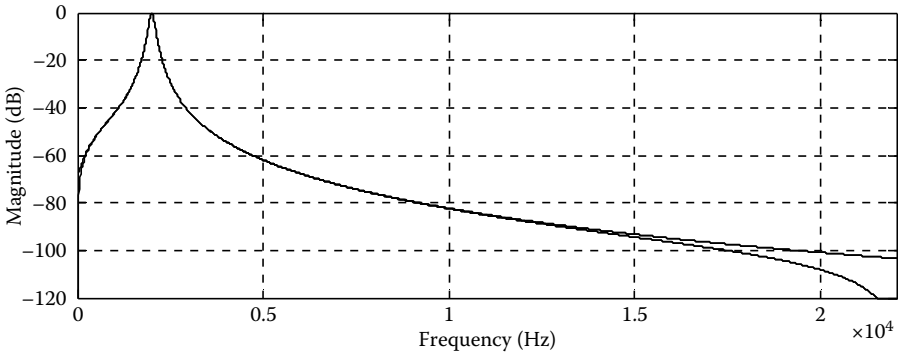


FIGURE 7.20 Magnitude of the FT and the DTFT, for $f_s = 44.1$ kHz.

TABLE 7.2 DTFT Properties

Linearity	$c_1x_1(nT)+c_2x_2(nT) \leftrightarrow c_1X_1(e^{j\omega T})+c_2X_2(e^{j\omega T})$
Time shift	$x(nT-n_0T) \leftrightarrow X(e^{j\omega T})e^{-j\omega n_0 T}$
Modulation	$x(nT)e^{j\omega_0 nT} \leftrightarrow X(e^{j(\omega-\omega_0)T})$
Linear convolution	$x_1(nT)^*x_2(nT)=\sum_{i=-\infty}^{+\infty} x_1((n-i)T)x_2(iT) \leftrightarrow X_1(e^{j\omega T})X_2(e^{j\omega T})$
Difference	$x(nT)-x((n-1)T) \leftrightarrow (1-e^{j\omega T})X(j\omega)$
Multiplication	$x_1(nT)x_2(nT) \leftrightarrow \frac{1}{2\pi}X_1(e^{j\omega T}) \otimes X_2(e^{j\omega T})$
Energy (Parseval's theorem)	$\sum_{n=-\infty}^{+\infty} \lvert x(nT)\rvert^2 = \frac{1}{2\pi}\int_0^{\omega_s} \lVert X(e^{j\omega T})\rVert^2 d\omega$

7.1.13 DTFT Properties

The DTFT has properties, most of which are analogous to properties of the FT. In Table 7.2 are given a few important properties of the DTFT.

7.1.14 Example and Discussion

For practical signals we only have a finite number of samples of $x(nT)$, say N samples, and therefore we must truncate the summation in Equation 7.37, and compute $X(e^{j\omega T})$ by digital means. This means that we can find the DTFT for only a finite number of frequencies. Let us evaluate the truncated DTFT at the frequencies given by

$$\omega = k\frac{\omega_s}{N}, \quad k = 0, 1, \ldots, N-1$$

and then the DTFT becomes

$$X(e^{j\omega T})\big|_{\omega=k\frac{\omega_s}{N}} = \sum_{n=0}^{N-1} x(nT)e^{-jk\frac{\omega_s}{N}nT} = \sum_{n=0}^{N-1} x(nT)e^{-j\frac{2\pi}{N}kn} = X(k) \tag{7.44}$$

which is the DFT of $x(nT)$, $n=0, 1, \ldots, N-1$.

Let us window the signal shown in Figure 7.19. Regardless of how wide the window is, the windowed signal $x_p(t)$ is one period of a periodic signal to which we are applying the DFT. Therefore, we should consider using a window such as $w_H(t)$ over the time range $0 \le t < T_0$. From Figure 7.19 we see that using $T_0 = 5.0$ ms does not incur discontinuities at period boundaries of the periodic extension of $x_p(t)$. Let us obtain $N = 226$ samples of $x(t)$ at the same rate $f_s = 44.1$ kHz. Then the duration of $x_p(t)$ becomes $T_0 = 5.1247$ ms. Figure 7.21 shows the magnitude of the FT of $x(t)$ and its approximation that was found with the DFT. While the DFT produces a peak near f_0, there is aliasing error above and below this frequency. This can be improved if $N(T_0)$ is increased.

7.1.15 Linear and Time Invariant Discrete-Time Systems

The discrete-time system (DTS) that we shall study is described by the block diagram shown in Figure 7.22, where $x(n)$ is the input to the system and $y(n)$ is the response or output of the system. We assume that the output and input are related by the difference equation given by

$$y(n)+a_1y(n-1)+\cdots+a_Ny(n-N)=b_0x(n)+b_1x(n-1)+\cdots+b_Mx(n-M) \tag{7.45}$$

where the a_i, $i=1,\ldots,N$ and b_i, $i=0,1,\ldots,M$ are real constants.

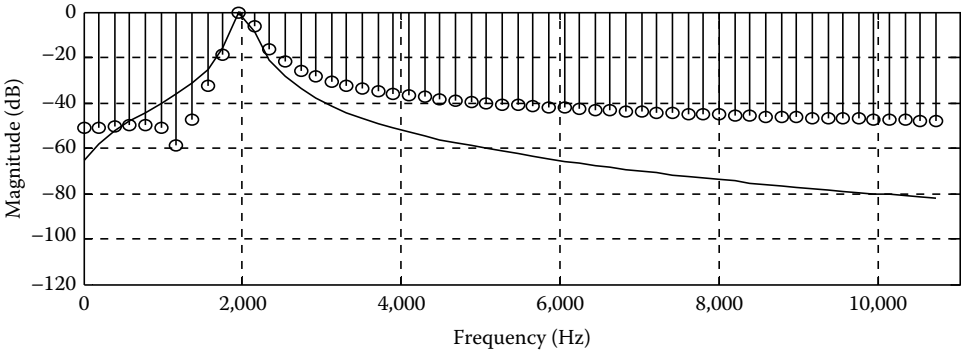


FIGURE 7.21 DFT approximation of the FT of $x(t)$.

To operate the DTS, reorganize Equation 7.45 to express the present output in terms of past outputs and present and past inputs, and we get

$$y(n) = \sum_{i=0}^{M} b_i x(n-i) - \sum_{i=1}^{N} a_i y(n-i) \qquad (7.46)$$

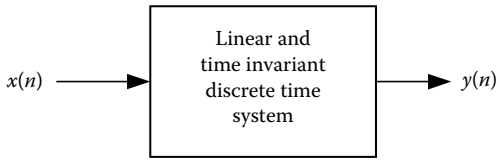


FIGURE 7.22 Block diagram of a DTS.

Suppose an input is applied starting at some time, say $n=0$. Then, to compute $y(n=0)$ given the first input $x(n=0)$, we must know

$$y(-1), y(-2), \ldots, y(-N)$$

which are the outputs prior to applying the input. These outputs are called the initial conditions of the DTS. Having computed $y(0)$, we can then compute $y(n=1)$, once we receive $x(n=1)$, and this can be continued for as long as we want to operate the DTS.

If in Equation 7.45 the right-hand side includes terms such as $b_{-1}x(n+1), b_{-2}x(n+2), \ldots$, then to compute a present output $y(n)$, knowledge of future inputs is required, and DTS cannot be operated in real time. Such a system is said to be noncausal. A system is said to be causal if a present output does not depend on a future value of the input. To operate a system in real time, it must be a causal system. However, when an entire history of an input is known, then at any time within this history, future values of the input are known, and a noncausal system can be operated. Both a causal and a noncausal system can be operated this way, which is referred to as off-line processing.

The DTS described by Equation 7.45 is a LTI system. Suppose all initial conditions are zero; we say the system is initially at rest. Let $y(n)=y_1(n)$ denote the response when the input is $x(n)=x_1(n)$, and similarly for $y_2(n)$ and $x_2(n)$. Then the DTS is a linear system if when the input is $x(n)=c_1x_1(n)+c_2x_2(n)$, then the response is given by $y(n)=c_1y_1(n)+c_2y_2(n)$ for any constants c_1 and c_2. The DTS is a time invariant system if when $y(n)=y_1(n)$ is the reponse to $x(n)=x_1(n)$, then $y(n)=y_2(n)=y_1(n-i)$ is the response to $x(n)=x_2(n)=x_1(n-i)$, for any time shift i.

A very useful function to know about an LTI DTS is its unit pulse response. This is the solution of Equation 7.45, which is denoted by $h(n)$, when $x(n)=\delta(n)$ and all initial conditions are zero. If the system is causal, then for $n<0$, we must have $h(n)=0$. To find $h(n)$, we first solve the simpler problem given by

$$v(n) + a_1 v(n-1) + \cdots + a_N v(n-N) = x(n) \qquad (7.47)$$

where initial conditions are zero and $x(n) = \delta(n)$. Therefore, we already have $v(0) = 1$. If we can determine the solution $v(n)$ of Equation 7.47, then, since the DTS is LTI and in view of the right-hand side of Equation 7.45, $h(n)$ is given by

$$h(n) = b_0 v(n) + b_1 v(n-1) + \cdots + b_M v(n-M) \tag{7.48}$$

For $n > 0$, Equation 7.47 becomes

$$v(n) + a_1 v(n-1) + \cdots + a_N v(n-N) = 0 \tag{7.49}$$

which is a homogeneous equation. Therefore, we can instead solve Equation 7.49 for initial conditions given by $v(0) = 1$, $v(-1) = 0, \ldots, v(N-1) = 0$. We want to find a function that can satisfy Equation 7.49. Let us try $v(n) = K\gamma^n$, for some nonzero constants K and γ. Substituting this $v(n)$ into Equation 7.49 gives

$$K\gamma^n + a_1 K\gamma^{n-1} + a_2 K\gamma^{n-2} + \cdots + a_N K\gamma^{n-N} = 0$$

and after multiplying by K^{-1} and γ^{-n} we get

$$Q(\gamma^{-1}) = 1 + a_1\gamma^{-1} + a_2\gamma^{-2} + \cdots + a_{N-1}\gamma^{-(N-1)} + a_N\gamma^{-N} = 0 \tag{7.50}$$

which is called the characteristic equation of the DTS, and $Q(\gamma^{-1})$ is called the characteristic polynomial. The characteristic equation has N roots, γ_i, $i = 1, 2, \ldots, N$, which can be real or complex, and Equation 7.50 can be factored to become

$$Q(\gamma^{-1}) = \prod_{i=1}^{N} (1 - \gamma_i\gamma^{-1}) = 0$$

Notice that these roots are not determined by the input. They are a characteristic of the system. Let us assume that they are distinct. Since each function γ_i^n can satisfy Equation 7.49, the solution is a linear combination of these exponential functions, and we have

$$v(n) = \sum_{i=1}^{N} K_i\gamma_i^n$$

The N constants K_i can be found by setting up N equations in N unknowns by applying the initial conditions of Equation 7.49. With $v(n)$ and Equation 7.48 we can then obtain the unit pulse response $h(n)$.

It is important to note that the unit pulse response is a linear combination of exponential functions with behavior that is determined by the roots of the characteristic equation. These roots can be real or they can occur in complex conjugate pairs. For example, if a root γ is real, then $h(n)$ will contain the term $K\gamma^n$ that may decrease exponentially to zero if $||\gamma|| < 1$, remain at unity if $\gamma = 1$, or increase exponentially if $||\gamma|| > 1$. Or, if a complex conjugate pair of roots occurs, then $h(n)$ will contain the sum of terms $K\gamma^n + K^*(\gamma^*)^n = 2||K||\ ||\gamma||^n \cos(n\angle\gamma + \angle K)$ that will oscillate and increase or decrease exponentially. If we require that $\lim_{n\to\infty} h(n) = 0$, then all roots of the characteristic equation must satisfy $||\gamma_i|| < 1$, $i = 1, 2, \ldots, N$.

A DTS can be categorized into two very different classes. Suppose that in Equation 7.45, $a_i = 0$ and $i = 1, 2, \ldots, N$. Then, Equation 7.46 becomes

$$y(n) = \sum_{i=0}^{M} b_i x(n-i) \tag{7.51}$$

and the present output depends on present and past inputs. If the input has a finite duration, then the response will have a finite duration. In fact, if $x(n) = \delta(n)$, then the unit pulse response is given by

$$h(n) = \sum_{i=0}^{M} b_i \delta(n - i)$$

which has a finite duration. Therefore, the DTS described by Equation 7.51 is called an FIR (finite impulse response) system, because the unit pulse response has a finite duration. However, if at least one a_i in Equation 7.45 is not zero, then in Equation 7.46 the present output will depend on the past outputs, which means there is feedback in the system. In this case the response to $x(n) = \delta(n)$ can never become and remain zero indefinitely. Such a DTS is called an IIR (infinite impulse response) system, because the response to a finite duration input will have an infinite duration.

7.1.16 Convolution

We can find the response of LTI DTS initially at rest to any input by writing the input as in Equation 7.8. Since the DTS is LTI, the response to $x(i)\delta(n - i)$ is $x(i)h(n - i)$, and the response to $x(n)$ is given by

$$y(n) = \sum_{i=-\infty}^{+\infty} h(n - i)x(i) = h(n)^* x(n) = \sum_{i=-\infty}^{+\infty} h(i)x(n - i) \tag{7.52}$$

which is called the linear convolution operation. Since the DTS is time invariant, the response to $x(n - n_0)$ is

$$y(n - n_0) = h(n)^* x(n - n_0)$$

7.1.17 Stability

A DTS is said to be bounded-input-bounded-output (BIBO) stable if for a bounded input the output is bounded. If the input is bounded, then for some real and positive number $x_{\max}$, we have $||x(n)|| \leq x_{\max}$ for all n. For a bounded output, Equation 7.52 becomes

$$\|y(n)\| = \left\| \sum_{i=-\infty}^{+\infty} h(i)x(n - i) \right\| \leq \sum_{i=-\infty}^{+\infty} \|h(i)x(n - i)\| \leq \sum_{i=-\infty}^{+\infty} \|h(i)\| x_{\max} = x_{\max} \sum_{i=-\infty}^{+\infty} \|h(i)\| < \infty$$

A sufficient condition for BIBO stability of an LTI DTS is that the unit pulse response must be absolutely summable.

Recall that the nature of the unit pulse response is determined by the roots of the characteristic equation (Equation 7.50). If all roots satisfy the condition $||\gamma_i|| < 1$, then we can show that $h(n)$ is absolutely summable. Moreover, we then have $\lim_{n\to\infty} h(n) = 0$. This means that if we apply a finite duration input to an initially at rest system, then eventually, the system will approach the rest state. An important distinction between FIR and IIR systems is that an FIR system is unconditionally stable, while an IIR system can be stable or unstable, depending on the roots of the characteristic equation.

7.1.18 Frequency Response

A major concern is the steady-state response of the DTS described by Equation 7.46 when the input is a sinusoid given by

$$x(n) = A \cos(\omega n T + \phi)$$

If the DTS is stable, then, in steady state, the response will be a sinusoid. It will be easier to determine the steady-state response if we work with an input $v(n)$ defined by

$$v(n) = e^{j\omega nT}$$

and let $w(n)$ denote the response to $v(n)$. Since

$$x(n) = \frac{A}{2}e^{j\phi}v(n) + \frac{A}{2}e^{-j\phi}v^{*}(n)$$

and since the DTS is a linear system, then the response $y(n)$ to $x(n)$ must be

$$y(n) = \frac{A}{2}e^{j\phi}w(n) + \frac{A}{2}e^{-j\phi}w^{*}(n)$$

With Equation 7.52, $w(n)$ is given by

$$\begin{aligned} w(n) &= \sum_{i=-\infty}^{+\infty} h(i)v(n-i) = \sum_{i=-\infty}^{+\infty} h(i)e^{j\omega(n-i)T} = e^{j\omega nT}\sum_{i=-\infty}^{+\infty} h(i)e^{-j\omega iT} \\ &= H(e^{j\omega T})e^{j\omega nT} \end{aligned}$$

where $H(e^{j\omega T})$ is the DTFT of the unit pulse response $h(n)$, and therefore $y(n)$ becomes

$$y(n) = \frac{A}{2}e^{j\phi}H(e^{j\omega T})e^{j\omega nT} + \frac{A}{2}e^{-j\phi}H^{*}(e^{j\omega T})e^{-j\omega nT}$$

Expressing $H(e^{j\omega T})$ in polar form, then with Euler's identity we get

$$y(n) = A\left|H(e^{j\omega T})\right|\cos\left(\omega nT + \phi + \angle H(e^{j\omega T})\right) \tag{7.53}$$

Let us consider another approach to find $y(n)$. First, let us substitute $v(n)$ into Equation 7.45 to get

$$\begin{aligned} w(n) + a_1 w(n-1) + \cdots + a_N w(n-N) &= b_0 e^{j\omega nT} + b_1 e^{j\omega(n-1)T} + \cdots + b_M e^{j\omega(n-M)T} \\ &= (b_0 + b_1 e^{-j\omega T} + \cdots + b_M e^{-j\omega MT})e^{j\omega nT} \end{aligned}$$

Since the right-hand side of this equation is an exponential function, then the left-hand side must also be an exponential function. Let us try $w(n) = We^{j\omega nT}$, for some unknown constant W, and then we get

$$W(1 + a_1 e^{-j\omega T} + \cdots + a_N e^{-j\omega NT})e^{j\omega nT} = (b_0 + b_1 e^{-j\omega T} + \cdots + b_M e^{-j\omega MT})e^{j\omega nT}$$

Therefore, we can solve for W, and $w(n)$ is

$$w(n) = \frac{(b_0 + b_1 e^{-j\omega T} + \cdots + b_M e^{-j\omega MT})}{(1 + a_1 e^{-j\omega T} + \cdots + a_N e^{-j\omega NT})}e^{j\omega nT}$$

Then, $H(e^{j\omega T})$ is given by

$$H(e^{j\omega T}) = \frac{(b_0 + b_1 e^{-j\omega T} + \cdots + b_M e^{-j\omega MT})}{(1 + a_1 e^{-j\omega T} + \cdots + a_N e^{-j\omega NT})} \tag{7.54}$$

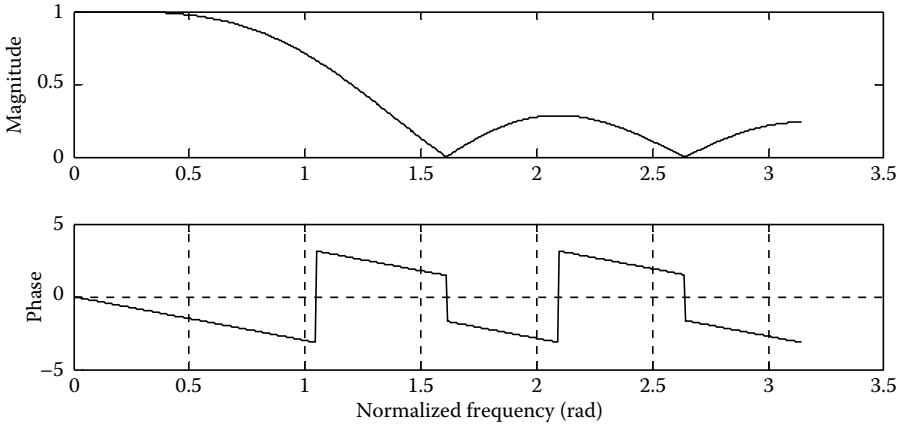


FIGURE 7.23 Magnitude and phase frequency response of an FIR filter.

which we can find by inspection of the given difference equation. In view of Equation 7.53 we have that $H(e^{j\omega T})$ gives the frequency response of the DTS. Notice that $H^*(e^{j\omega T}) = H(e^{-j\omega T})$, and therefore $||H(e^{j\omega T})||$ is an even periodic function of ω and $\angle H(e^{j\omega T})$ is an odd periodic function of ω.

7.1.19 Examples and Discussion

With Equation 7.54 we can assess the frequency selective behavior of an LTI DTS. Consider the FIR DTS given by

$$y(n) = 1/21[-2x(n) + 3x(n-1) + 6x(n-2) + 7x(n-3) + 6x(n-4) + 3x(n-5) - 2x(n-6)]$$

The frequency response is

$$H(\mathrm{e}^{j\omega T}) = 1/21[-2 + 3\mathrm{e}^{-j\omega T} + 6\mathrm{e}^{-j2\omega T} + 7\mathrm{e}^{-j3\omega T} + 6\mathrm{e}^{-j4\omega T} + 3\mathrm{e}^{-j5\omega T} - 2\mathrm{e}^{-j6\omega T}]$$

To see the frequency selective behavior of this DTS we could plot the magnitude of the frequency response for $0 \leq \omega \leq \omega_s/2$. Instead, let us plot the magnitude versus $\theta = \omega T$, where $0 \leq \theta \leq \pi$. Figure 7.23 shows the magnitude and phase frequency response, and we see that this DTS is a low-pass filter. Furthermore, this FIR filter has a phaseshift characteristic that varies linearly with frequency. The step changes in phase are due to wraparound when the phase changes from, for example, $-\pi$ to $+\pi$.

As a comparison, let us assess the frequency selective behavior of the IIR LTI DTS described by

$$y(n) = 0.09780\ x(n) + 0.19560\ x(n-1) + 0.09780\ x(n-2) + 0.94175\ y(n-1) - 0.33296\ y(n-2)$$

with $H(\mathrm{e}^{j\omega T})$ given by

$$H(\mathrm{e}^{j\omega T}) = \frac{0.09780 + 0.19560\mathrm{e}^{-j\omega T} + 0.09780\mathrm{e}^{-j2\omega T}}{1 - 0.94175\mathrm{e}^{-j\omega T} + 0.33296\mathrm{e}^{-j2\omega T}}$$

Figure 7.24 shows magnitude and phase frequency response of this IIR DTS. This is a low-pass filter, and the transition from the passband to the stopband is not very sharp. The phaseshift does not vary linearly with frequency.

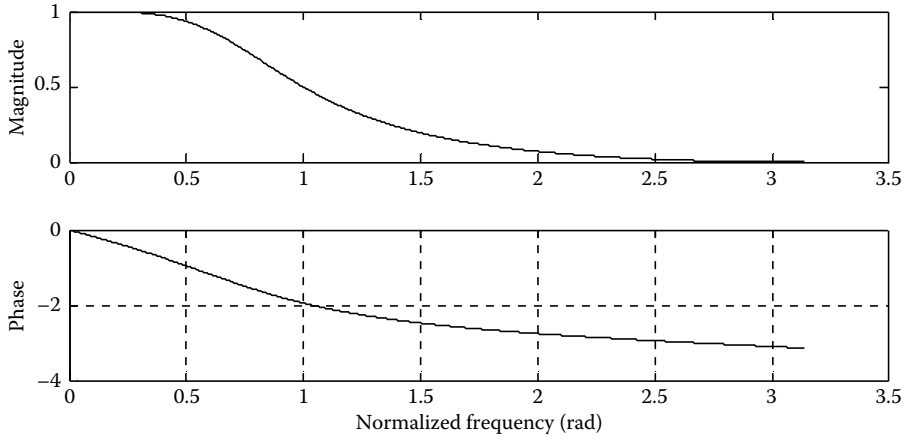


FIGURE 7.24 Magnitude and phase frequency response of an IIR filter.

These two filters have similar performance. Each has advantages and disadvantages compared to the other filter. For example, the FIR filter is unconditionally stable and it has a linear phase characteristic, and the IIR filter has a sharper magnitude frequency response with less computation than required by the FIR filter.

7.1.20 Ideal Digital Filters

We can design the coefficients of the digital filter described by Equation 7.45 to achieve filter performance that comes arbitrarily close to an ideal filter performance. Figure 7.25 shows the ideal performance of several standard filter types. The frequency response for each filter is specified over the frequency range, $-\omega_s/2 \leq \omega \leq \omega_s/2$, which is the frequency range of the filter input. Then, this frequency response is extended periodically to integer multiples of ω_s.

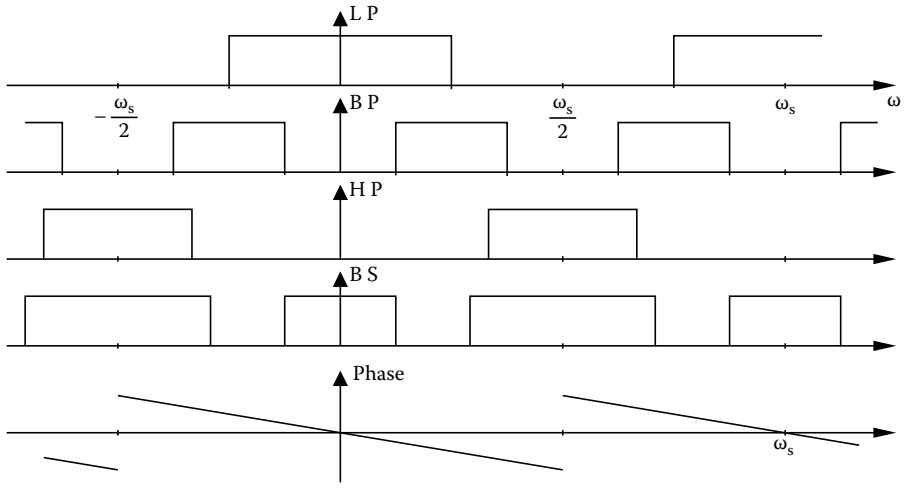


FIGURE 7.25 Ideal frequency response of a low-pass (LP), band-pass (BP), high-pass (HP), and band-stop (BS) digital filter.

To justify an ideal linear phase characteristic, let us consider an input $x(t) = A_1 \cos(\omega_1 t + \phi_1) + A_2 \cos(\omega_2 t + \phi_2)$ of an analog filter, where ω_1 and ω_2 are within the passband of the analog filter. To preserve the input wave shape, the output $y(t)$ can be a delayed version of the input. Now we have

$$\begin{aligned} y(t) = x(t - t_0) &= A_1 \cos(\omega_1(t - t_0) + \phi_1) + A_2 \cos(\omega_2(t - t_0) + \phi_2) \\ &= A_1 \cos(\omega_1 t + \phi_1 - t_0\omega_1) + A_2 \cos(\omega_2 t + \phi_2 - t_0\omega_2) \end{aligned}$$

and to preserve the input wave shape, the input to output phase change must be proportional to frequency.

7.1.21 *z*-Transform

While the DTFT is useful for spectral analysis, its wider application to system analysis is limited, because, for example, not even the DTFT of a discrete-time sinusoidal signal exists. Given a discrete-time signal $x(nT)$, let us consider instead the DTFT given by

$$Y(\sigma, e^{j\omega T}) = \text{DTFT}\{e^{-\sigma nT} x(nT)\} = \sum_{n=-\infty}^{+\infty} e^{-\sigma nT} x(nT) e^{-j\omega nT} \tag{7.55}$$

for some real number σ. Assuming that $Y(\sigma, e^{j\omega T})$ exists, taking its IDTFT gives

$$e^{-\sigma nT} x(nT) = \frac{1}{\omega_s} \int_{-\frac{\omega_s}{2}}^{+\frac{\omega_s}{2}} Y(\sigma, e^{j\omega T}) e^{j\omega nT} d\omega$$

and after multiplying both sides of this equation by $e^{\sigma nT}$ we get

$$x(nT) = \frac{1}{\omega_s} \int_{-\frac{\omega_s}{2}}^{+\frac{\omega_s}{2}} e^{\sigma nT} Y(\sigma, e^{j\omega T}) e^{j\omega nT} d\omega \tag{7.56}$$

For convenience, let

$$z = e^{\sigma T} e^{j\omega T} = e^{(\sigma + j\omega)T} = \|z\| e^{j\angle z} \tag{7.57}$$

and let $X(z)$ denote $Y(\sigma, e^{j\omega T})$, and therefore, Equation 7.55 becomes

$$X(z) = \sum_{n=-\infty}^{+\infty} x(n) z^{-n} \tag{7.58}$$

which is called the bilateral z-transform (BZT) of $x(n)$. We write $X(z) = Z\{x(n)\}$. In Equation 7.56, let us change the integration variable from the real variable ω to the complex variable z. When $\omega = -\omega_s/2$, $z = e^{\sigma T} e^{-j\pi}$, and when $\omega = +\omega_s/2$, $z = e^{\sigma T} e^{j\pi}$, and therefore z follows a circle of radius $||z|| = e^{\sigma T}$. We also have

$$\frac{dz}{d\omega} = jT e^{\sigma T} e^{j\omega T} = jTz$$

and therefore Equation 7.56 becomes

$$x(nT) = \frac{1}{\omega_s} \oint X(z) z^n \frac{dz}{jTz} = \frac{1}{j2\pi} \oint X(z) z^{n-1} dz \tag{7.59}$$

which is called the inverse z-transform, denoted by $x(n) = Z^{-1}\{X(z)\}$. Notice that the integration path is a circle with a radius determined by σ.

7.1.22 Bilateral z-Transform Properties

The BZT has many useful properties. For our development we will use the properties given in Table 7.3.

7.1.23 z-Plane

To see the utility of introducing the factor $e^{-\sigma nT}$ in Equation 7.55, let us obtain the BZT of $x(n) = \gamma^n u(n)$ for any real or complex number γ. With Equation 7.58 we have

$$X(z) = \sum_{n=-\infty}^{+\infty} \gamma^n u(n) z^{-n} = \sum_{n=0}^{+\infty} \left(\frac{\gamma}{z}\right)^n = \lim_{N\to\infty} \sum_{n=0}^{N-1} \left(\frac{\gamma}{z}\right)^n = \lim_{N\to\infty} \frac{1 - \left(\frac{\gamma}{z}\right)^N}{1 - \left(\frac{\gamma}{z}\right)} = \frac{1}{1 - \gamma z^{-1}}, \quad \left\|\frac{\gamma}{z}\right\| < 1$$

If $||\gamma|| > 1$, then $x(n)$ increases exponentially, and it does not have a DTFT. However, $x(n)$ does have a BZT, because we restrict z with σ to satisfy the convergence condition, $||z|| = e^{\sigma T} > ||\gamma||$, which requires $\sigma > \frac{1}{T} \ln\|\gamma\|$. If σ satisfies this condition, then the product $e^{-n\sigma T}\ \gamma^n$ decreases exponentially. The function $e^{-n\sigma T}$ is called the convergence factor.

We say that $x(n)$ and $X(z)$ are a BZT pair, denoted by

$$\gamma^n u(n) \leftrightarrow 1/(1 - \gamma z^{-1}), \quad \|z\| > \|\gamma\| \tag{7.60}$$

This transform pair includes signals such as

$$x(n) = u(n), \quad \gamma = 1$$
$$x(n) = \cos(\theta n) u(n) = \frac{1}{2} e^{j\theta n} u(n) + \frac{1}{2} e^{-j\theta n} u(n), \quad \gamma = e^{j\theta} \quad \text{and} \quad \gamma = e^{-j\theta}$$

and many other signals by applying BZT properties. For example, notice that Equation 7.60 can be obtained with $Z\{u(n)\}$ and the modulation property.

If $||\gamma|| < 1$, then $x(n)$ decreases exponentially, and it has a DTFT. It also has the BZT given in Equation 7.60. With $||\gamma|| < 1$, the condition $\sigma > \frac{1}{T} \ln\|\gamma\|$ allows σ to be negative, and with a negative σ the convergence factor increases exponentially, while the product $e^{-n\sigma T}\gamma^n$ still decreases exponentially. We can even set σ to $\sigma = 0$, which with Equation 7.57 converts the BZT to the DTFT.

TABLE 7.3 BZT Properties

Linearity	$c_1 x_1(n) + c_2 x_2(n) \leftrightarrow c_1 X_1(z) + c_2 X_2(z)$
Time shift	$x(n - n_0) \leftrightarrow X(z) z^{-n_0}$
Modulation	$x(n)\alpha^n \leftrightarrow X(z\alpha^{-1})$
Linear convolution	$x_1(n) * x_2(n) = \sum_{i=-\infty}^{+\infty} x_1(n-i) x_2(i) \leftrightarrow X_1(z) X_2(z)$

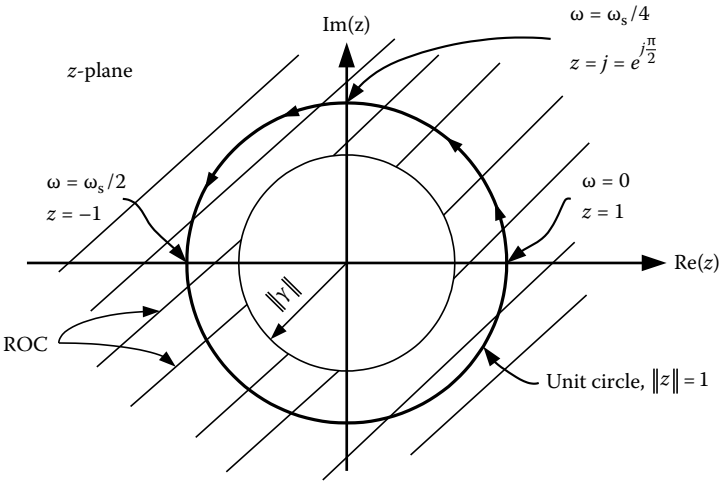


FIGURE 7.26 The z-plane, showing the unit circle and an ROC.

These different situations can be nicely pulled together by describing them with respect to the unit circle in the z-plane, as shown in Figure 7.26. The convergence condition, $||z|| > ||\gamma||$, does not impose any restriction on the angle of z. Therefore, the BZT of $x(n) = \gamma^n u(n)$ converges for all z outside a circle in the z-plane having a radius equal to $||\gamma||$, and this region of the z-plane is called the region of convergence (ROC). If $||\gamma|| < 1$, which means that $x(n)$ has a DTFT, then the ROC is outside a circle that is inside the unit circle in the z-plane (see Figure 7.26). However, if $||\gamma|| > 1$, which means that the DTFT of $x(n)$ does not exist, then the ROC is outside a circle that is outside the unit circle. This means that if the ROC of the BZT includes the unit circle, then $X(z)$ can be evaluated for $z = e^{j\omega T}$ ($||z|| = 1$), and a BZT can be converted to a DTFT by setting z to $z = e^{j\omega T}$.

The ROC is an important information about the BZT of a signal, because there can be an ambiguity between $X(z)$ and its inverse z-transform. To see why this is true, let us find the BZT of $x(n) = -\gamma^n u(-n-1)$, which is zero for $n \geq 0$. Applying Equation 7.58 gives

$$X(z) = \sum_{n=-\infty}^{+\infty} -\gamma^n u(-n-1) z^{-n} = -\sum_{n=-\infty}^{-1} \gamma^n z^{-n} = -\sum_{n=1}^{+\infty} (\gamma^{-1} z)^n = -\sum_{n=0}^{+\infty} (\gamma^{-1} z)^n + 1$$
$$= 1 - \frac{1}{1-\gamma^{-1} z} = \frac{1}{1-\gamma z^{-1}}, \left\|\gamma^{-1} z\right\| < 1 \rightarrow \|z\| < \|\gamma\| \tag{7.61}$$

Here the ROC is the interior of a circle in the z-plane. If $||\gamma|| > 1$, then the ROC includes the unit circle and $x(n)$ decreases exponentially as $n \rightarrow -\infty$, which means that the DTFT of $x(n)$ exists. However, if $||\gamma|| < 1$, then the ROC does not include the unit circle and $x(n)$ increases exponentially as $n \rightarrow -\infty$, which means that the DTFT of $x(n)$ does not exist.

If the BZT of Equation 7.60 or 7.61 is given, we cannot know the inverse z-transform unless we also know the convergence condition. In Equation 7.60, $x(n)$ is said to be a right-handed signal, and its ROC is a region in the z-plane outside a circle, and in Equation 7.61, $x(n)$ is said to be a left-handed signal, and its ROC is a region in the z-plane inside a circle. Consider how the ROC must change if we time reverse a signal. For either kind of signal, right-handed or left-handed, if the ROC includes the unit circle, then it has a DTFT. We will work with right-handed signals unless explicitly stated otherwise, in which case the ROC is the type given in Equation 7.60 and illustrated in Figure 7.26.

7.1.24 Transfer Function

Given an input $x(n)$ of an LTI DTS, the convolution Equation 7.52 of the input with the unit pulse response $h(n)$ gives the response of the DTS under zero initial conditions. Applying the convolution property in Table 7.3 to Equation 7.52 gives

$$Y(z) = H(z)X(z) \tag{7.62}$$

where $H(z) = Z\{h(n)\}$, which is called the transfer function of the system. We assume that the DTS is causal, and therefore $h(n)$ is a right-handed signal. If $x(n) = \delta(n)$, then $X(z) = 1$, and, as expected, $y(n) = h(n)$.

Let us apply the linearity and time-shift properties of the BZT, and take the BZT term-by-term of Equation 7.45 to get

$$Y(z) + a_1 Y(z)z^{-1} + \cdots + a_N Y(z)z^{-N} = b_0 X(z) + b_1 X(z)z^{-1} + \cdots + b_M X(z)z^{-M}$$

Having converted the difference equation to an algebraic equation permits us to find

$$Y(z) = \frac{b_0 + b_1 z^{-1} + \cdots + b_M z^{-M}}{1 + a_1 z^{-1} + \cdots + a_N z^{-N}} X(z) \tag{7.63}$$

and therefore, $H(z)$ in Equation 7.62 is given by

$$H(z) = \frac{b_0 + b_1 z^{-1} + \cdots + b_M z^{-M}}{1 + a_1 z^{-1} + \cdots + a_N z^{-N}} = \frac{P(z^{-1})}{Q(z^{-1})} = \frac{b_0 \prod_{i=1}^{M} (1 - z_i z^{-1})}{\prod_{i=1}^{N} (1 - p_i z^{-1})} \tag{7.64}$$

The z_i, $i = 1, \ldots, M$ are called the zeros of $H(z)$, and the p_i, $i = 1, \ldots, N$ are called the poles of $H(z)$.

Comparing Equations 7.64 and 7.50, we see that the denominator of $H(z)$ is the characteristic polynomial, and therefore the poles of $H(z)$ are the roots of the characteristic equation, which determine the behavior of the unit pulse response. Therefore, $h(n)$ is a linear combination of the exponential functions, p_i^n, $i = 1, 2, \ldots, N$. This means that the DTS is stable if all the poles of $H(z)$ are inside the unit circle in the z-plane. Furthermore, the ROC of $H(z) = Z\{h(n)\}$ is outside a circle with a radius equal to the magnitude of the pole having the largest magnitude among all the poles of $H(z)$. Therefore, if the DTS is stable, $H(z)$ can be converted to the DTFT with $z = e^{j\omega T}$.

Comparing Equations 7.64 and 7.54, we see that the frequency response of the DTS can be found by evaluating $H(z)$ on the unit circle, where $z = e^{j\omega T}$. This is depicted in Figure 7.26, where for $\omega = 0$, $z = 1$. As ω is increased from $\omega = 0$ to $\omega = \omega_s/2$, the angle of z, $\angle z = \omega T$, changes from $\angle z = 0$, where $z = 1$, to $\angle z = \pi$, where $z = -1$.

7.1.25 Unilateral z-Transform

Consider a signal $x(n)$, which is defined over $-\infty < n < +\infty$. Generally, $x(n)$ is not a right-handed or a left-handed signal. Let us express $x(n)$ as

$$x(n) = x_l(n) + x_r(n) = x(n)u(-(n+1)) + x(n)u(n)$$

where $x_l(n) = 0$, $n \geq 0$ and $x_r(n) = 0$, $n < 0$, so as to have $x_l(n)$ account for the left-handed part of $x(n)$, while $x_r(n)$ accounts for the right-handed part of $x(n)$. The BZT of $x(n)$ is

$$X(z) = \sum_{n=-\infty}^{-1} x(n)z^{-n} + \sum_{n=0}^{+\infty} x(n)z^{-n} = Z\{x_l(n)\} + Z\{x_r(n)\} = X_l(z) + X_r(z) \tag{7.65}$$

Suppose we are only interested to observe a signal $x(n)$ for $n \geq 0$, even if it is not strictly a right-handed signal. Then, with

$$Z^{-1}\{X(z)\} = x(n)$$

we can find $x(n)$ for $n \geq 0$, and we also get $x(n)$ for $n < 0$. However, we can also find $x(n)$ for $n \geq 0$ with

$$Z^{-1}\{X_r(z)\} = x_r(n)$$

Therefore, we will now only work with

$$X(z) = \sum_{n=0}^{+\infty} x(n)z^{-n} \tag{7.66}$$

which is called the unilateral z-transform, or just z-transform (ZT). This does not imply that $x(n)$ for $n < 0$ must be zero. Instead, we prefer to avoid involving $X_l(z)$.

The bilateral and unilateral z-transforms have many properties in common, while for some properties the difference is very useful. Consider the ZT of $y(n) = x(n-1)$. Notice that $y(0) = x(-1)$, and therefore

$$Z\{y(n) = x(n-1)\} = \sum_{n=0}^{+\infty} y(n)z^{-n} = \sum_{n=0}^{+\infty} x(n-1)z^{-n} = x(-1) + z^{-1}\sum_{n=0}^{+\infty} x(n)z^{-n}$$

Generally, we have

$$Z\{x(n-i)\} = x(-i) + z^{-1}Z\{x(n-(i-1))\}, \quad i \geq 1 \tag{7.67}$$

which is not the same as the time-shift property given in Table 7.3 for the BZT.

Unlike BZT, ZT can be used to find the complete response of an LTI DTS as described by Equation 7.45. We assume that the input $x(n)$ is zero for $n < 0$. Taking the term-by-term ZT of Equation 7.45 gives

$$\begin{aligned} Y(z) + a_1\big(y(-1) + z^{-1}Y(z)\big) + \cdots + a_N\big(y(-N) + z^{-1}Z\{y(n-(N-1))\}\big) \\ = b_0X(z) + b_1z^{-1}X(z) + \cdots + b_Mz^{-M}X(z) \end{aligned}$$

which can be rearranged to become

$$(1 + a_1z^{-1} + a_2z^{-2} + \cdots + a_Nz^{-N})Y(z) - Y_{IC}(z^{-1}) = (b_0 + b_1z^{-1} + b_2z^{-2} + \cdots + b_Mz^{-M})X(z)$$

where $Y_{IC}(z^{-1})$ accounts for all initial condition terms. If all initial conditions are zero, then $Y_{IC}(z^{-1}) = 0$. Solving for $Y(z)$ results in

$$Y(z) = \frac{Y_{\mathrm{IC}}(z^{-1})}{Q(z^{-1})} + H(z)X(z) \tag{7.68}$$

We can factor the denominator of $X(z)$, and write

$$X(z) = \frac{P_x(z^{-1})}{\prod_{i=1}^{N_x}(1 - p_{x,i}z^{-1})}$$

where the $p_{x,i}$, $i = 1, \ldots, N_x$, are the N_x poles of $X(z)$. Therefore, Equation 7.68 becomes

$$Y(z) = \frac{Y_{\mathrm{IC}}(z^{-1})}{\prod_{i=1}^{N}(1 - p_i z^{-1})} + \frac{P(z^{-1})}{\prod_{i=1}^{N}(1 - p_i z^{-1})} \frac{P_x(z^{-1})}{\prod_{i=1}^{N_x}(1 - p_x z^{-1})}$$

Let us assume that the poles of $H(z)$ and the poles $X(z)$ are distinct, and then a partial fraction expansion of the right-hand side of this equation has the form

$$Y(z) = \sum_{i=1}^{N} \frac{K_{\mathrm{IC},i}}{1 - p_i z^{-1}} + \sum_{i=1}^{N} \frac{K_i}{1 - p_i z^{-1}} + \sum_{i=1}^{N_x} \frac{K_{x,i}}{1 - p_{x,i} z^{-1}}$$

The complete solution $y(n)$ of Equation 7.45 is given by

$$y(n) = \sum_{i=1}^{N} K_{\mathrm{IC},i}(p_i)^n u(n) + \sum_{i=1}^{N} K_i(p_i)^n u(n) + \sum_{i=1}^{N_x} K_{x,i}(p_{x,i})^n u(n) \tag{7.69}$$

If the input $x(n)$ is zero, then $y(n)$ in Equation 7.69 consists of only the first sum term, which is called the zero-input response. If all initial conditions are zero, then $y(n)$ in Equation 7.69 consists of the second and third sum terms, which together are called the zero-state response. Furthermore, the first and second sum terms are due to the poles of $H(z)$, and together they are called the transient response. There can be a transient response even if all initial conditions are zero.

The third sum term in Equation 7.69 is a discrete-time function like the input, and it is called the steady-state response. In digital filtering a signal, this is the part of the complete response that is of main concern. We prefer that the transient response quickly decays to negligible levels. If the input is, for example, a single sinusoid, then $N_x = 2$, and the third sum term in Equation 7.69 will consist of the sum of two complex conjugate exponential functions that combine via Euler's identity to become a sinusoid, as given in Equation 7.53.

7.1.26 Conclusion

The concept of a FS was developed as the minimum mean square error solution for the problem of approximating a continuous-time periodic signal with a linear combination of sinusoidal functions. If the continuous-time periodic signal is everywhere continuous, then the approximation error is zero. By sampling FS, DFT was obtained, and its relationship to the spectral nature of the continuous-time signal was found. Depending on the bandwidth of the continuous-time signal, the sampling rate and the sampling duration, we came to understand the kinds of errors that can occur when the DFT is used to study the spectral nature of a signal. The DFT is a widely applied method of signal and system analysis, because the DFT can be efficiently computed with the FFT, which was demonstrated by the development of the decimation in time radix-2 FFT.

The FS concept was extended to aperiodic continuous-time signals, resulting in the FT. By sampling the inverse FT, the DTFT was obtained, and its relationship to the spectral nature of the continuous-time signal was found. The sampling theorem came by studying this relationship. The DTFT was applied to signal and system analysis, and for practical reasons, the DFT again played an important role.

To broaden the class of signals to which the DTFT can be applied, a convergence factor was introduced, resulting in the BZT. The relationship between the DTFT and BZT was thoroughly investigated to understand the role of the BZT in spectral analysis. The BZT was modified to work with right-handed signals resulting in the unilateral z-transform, or just ZT. The ZT was applied to find the complete response of a DTS, a main advantage of the ZT.

The next step for the reader may be to investigate methods for the design of digital filters, as can be found in the references and elsewhere, where the DFT, DTFT and ZT, which are all based on the concept of a FS, play major roles.

References

Ambardar, A., 2007, *Digital Signal Processing: A Modern Introduction*, Thomson, Publishing, Washington, DC.

Antoniou, A., 1993, *Digital Filters*, McGraw-Hill, New York.

Bracewell, R.N., 1978, *The Fourier Transform and Its Applications*, McGraw-Hill, New York.

McClellan, J.H., Schafer, R.N., and Yoder, M.A., 1997, *DSP First: A Multimedia Approach*, Prentice-Hall, Upper Saddle River, NJ.

Oppenheim, A.V. and Schafer, R.W., 1975, *Discrete Time Signal Processing*, Prentice-Hall, Upper Saddle River, NJ.

Orfanidis, S.J., 1996, *Introduction to Signal Processing*, Prentice-Hall, Upper Saddle River, NJ.

Priemer, R., 1991, *Introductory Digital Signal Processing*, World Scientific, Singapore.

Proakis, J.G. and Manolakis, D.G., 1996, *Digital Signal Processing: Principles, Algorithms and Applications*, Prentice-Hall, Upper Saddle River, NJ.

Rabiner, L.R. and Gold, B., 1974, *Theory and Application of Digital Signal Processing*, Prentice-Hall, Upper Saddle River, NJ.

Tan, T., 2008, *Digital Signal Processing, Fundamentals and Applications*, Academic Press, New York.

8

Digital Circuits

John P. Uyemura
Georgia Institute of Technology

Robert C. Chang
University of Southern California

Bing J. Sheu
Taiwan Semiconductor Manufacturing Company

8.1 MOS Logic Circuits **8**-1
Introduction • MOSFET Models for Digital Circuits • The Digital Inverter • nMOS Logic Gates • CMOS Inverter • Static CMOS Logic Gates • Dynamic CMOS Logic Gates
References **8**-15
8.2 Transmission Gates **8**-15
Digital Processing • Analog Processing
References **8**-33

8.1 MOS Logic Circuits

John P. Uyemura

8.1.1 Introduction

MOS-based technology has become the default standard for high-density logic designs for several reasons. The most obvious is that MOSFETs can be made with side dimensions of <0.1 μm (10^{-7} m), allowing for complex logic functions to be constructed in small areas. The section is an investigation of the basics of designing and characterizing logic gates in an MOS technology.

8.1.2 MOSFET Models for Digital Circuits

The properties of digital logic gates are derived from a large-signal analysis of the circuits. Because transistor characteristics are intrinsically nonlinear, accurate analytic modeling becomes quite complicated, and closed-form solutions can be difficult to come by. To overcome this problem, simplified MOSFET models are used to estimate the circuit operation in first-cut designs. Once the basic operation is established, computer simulations are used to obtain more accurate information.

Square-law models are useful for understanding the operation of MOS logic circuits. Consider first an *n*-channel, enhancement-mode MOSFET that has a threshold voltage $V_{Tn} > 0$. As shown in Figure 8.1, the primary device voltages are V_{DS}, V_{GS}, and V_{SB}. The value gate–source voltage V_{GS} relative to the threshold voltage V_{Tn} determines if drain current I_D flows. If $V_{GS}, < V_{Tn}$ then $I_D \approx 0$, establishing the condition of ''cutoff.'' Elevating the gate–source voltage to a value $V_{GS}, > V_{Tn}$ places the MOSFET into the ''active region'' where I_D will be nonzero if a drain–source voltage V_{DS} is applied; the value of I_D depends on the values of the device voltages.

To describe active operation, we introduce the drain–source saturation voltage $V_{DS,sat}$ defined by

$$V_{DS,sat} = V_{GS} - V_{Tn} \tag{8.1}$$

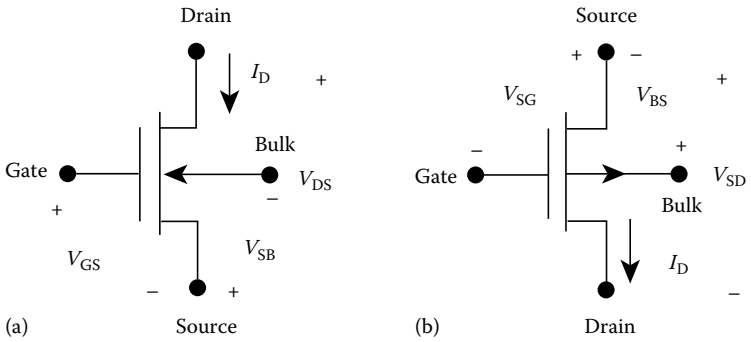


FIGURE 8.1 MOSFET symbols. (a) *n*-channel MOSFET; (b) *p*-channel MOSFET.

The threshold voltage is affected by the affected by the source–bulk (body) voltage V_{SB} by

$$V_{Tn} = V_{T0n} + \gamma\left(\sqrt{2|\phi_F| + V_{SB}} - \sqrt{2|\phi_F|}\right) \tag{8.2}$$

where

V_{T0n} is the nFET zero-body-bias threshold voltage
γ is the body bias (or, body effect) coefficient
ϕ_F is the bulk Fermi potential

When $V_{DS} \cdot V_{DS,sat}$, the MOSFET is nonsaturated with

$$I_D \approx \left(\frac{\beta_n}{2}\right)\left[2(V_{GS} - V_{Tn})V_{DS} - V_{DS}^2\right] \tag{8.3}$$

In this equation β_n is the device transconductance given by $\beta_n = k'_n(W/L)$, with k'_n the process transconductance in units of [A/V^2], W the channel width, and L the channel length; the width-to-length (W/L) is called the "aspect ratio" of the transistor. The process transconductance is given by $k'_n = \mu_n C_{ox}$ where μ_n is the electron surface mobility and C_{ox} the oxide capacitance per unit area. For an oxide layer with thickness t_{ox}, the MOS capacitance per unit area is calculated from

$$C_{ox} = \frac{\varepsilon_{ox}}{t_{ox}} \tag{8.4}$$

where ε_{ox} is the oxide permittivity. In the current technologies t_{ox} is smaller than about 60 Å. If $V_{DS} \geq V_{DS,sat}$, the MOSFET is saturated with

$$I_D \approx \left(\frac{\beta_n}{2}\right)(V_{GS} - V_{Tn})^2 \tag{8.5}$$

This ignores several effects, most notably that of "channel-length modulation," but is still a reasonable approximation for estimating basic performance parameters.

The structure of the MOSFET gives rise to several parasitic capacitances that tend to dominate the circuit performance. Two types of capacitors are contained in the basic model shown in Figure 8.2. The contributions C_{GS} and C_{GD} are due to the MOS layering of the gate–oxide semiconductor, which is the origin of the field effect. The total gate capacitance G_G is calculated from

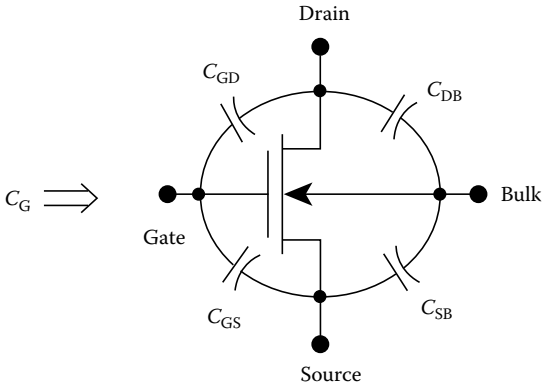


FIGURE 8.2 MOSFET capacitances.

$$C_G = C_{ox} WL \tag{8.6}$$

and the gate–source and gate–drain contributions can be approximated to first order by

$$C_{GD} \approx \frac{C_G}{2} \approx C_{GS} \tag{8.7}$$

Capacitors C_{DB} and C_{SB} are depletion contributions from the reverse-biased *pn* junction at the drain and source. These are nonlinear, voltage-dependent elements that decrease with increasing reverse voltage.

A *p*-channel MOSFET (pMOS or pFET) is the electrical complement of an *n*-channel device. An enhancement-mode pFET is defined to have a negative threshold voltage, i.e., $V_{Tp} < 0$. It is common to use device voltage of V_{SG}, V_{SD}, and V_{BS}, as shown in Figure 8.1 to describe the operation. Cutoff occurs if $V_{SG} < |V_{Tp}|$, while the device is active if $V_{SG} \geq |V_{Tp}|$. The saturation voltage of the pFET is defined by

$$V_{SD,sat} = V_{SG} - |V_{Tp}| \tag{8.8}$$

With $V_{SG} \geq |V_{Tp}|$ and $V_{SG} < V_{SD,sat}$, the transistor is nonsaturated with

$$I_D \approx \left(\frac{\beta_p}{2}\right)\left[2\left(V_{SG} - |V_{Tp}|\right)V_{SD} - V_{SD}^2\right] \tag{8.9}$$

For the pFET, β_p is the device transconductance $\beta_p = k'_p$ (W/L); where $k'_p = \mu_p C_{ox}$ is the process transconductance, and (W/L) is the aspect ratio of the device. In complementary metal–oxide–semiconductor (CMOS) inverters nFETs and pFETs are used in the same circuit, and it is important to note that $k'_n > k'_p$ due to the fact that the electron mobility is larger than the hole mobility, typically by a factor of 2 to 8.

It is often convenient to use the simplified MOSFET symbols shown in Figure 8.3. The polarity of the transistor (nMOS or pMOS) is made explicit by the absence or presence of the gate inversion "bubble," as shown. These symbols do not show the bulk electrode explicitly, but it is important to remember that all nFETs have their bulks connected to the lowest voltage in the circuit (usually ground), while all pFET bulks are connected to the highest voltage (usually the power supply V_{DD}).

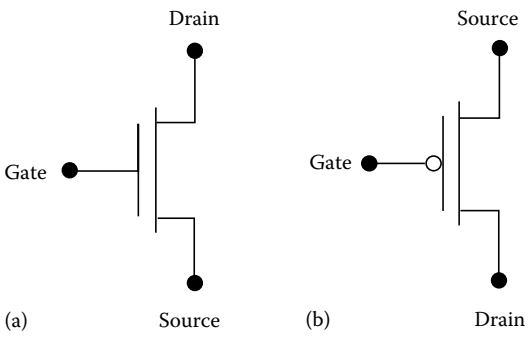


FIGURE 8.3 Simplified MOSFET symbols. (a) nMOSFET; (b) pMOSFET.

In digital circuit design, it is useful to model MOSFETs as voltage-controlled switches, as shown in Figure 8.4. The MOSFET switches are controlled by a gate input signal *G*, which is taken to be a Boolean variable. Employing a positive logic convention, $G = 0$ corresponds to a low voltage (below V_{Tn}), while $G = 1$ is a high voltage. The operation of the FET switches is straightforward. An input of $G = 0$ places the nFET into cutoff, corresponding to an OPEN switch; $G = 1$ implies active operation, and the switch is CLOSED. The pFET has a complementary behavior, with $G = 0$ giving a CLOSED switch, and $G = 1$ giving an OPEN switch.

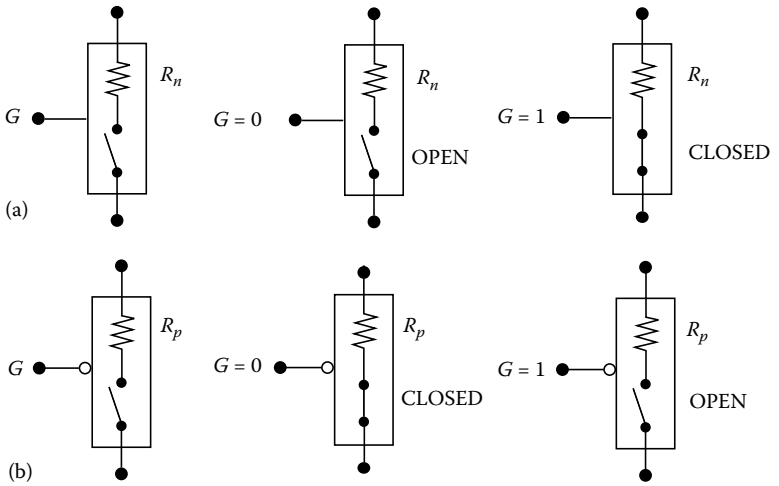


FIGURE 8.4 MOSFET switching models. (a) nFET switch model; (b) pFET switch model.

The switch models include parasitic drain-to-source resistance R_n and R_p, which are usually estimated using

$$R_n = \frac{1}{k'_n(\mathrm{W/L})_n(V_{\mathrm{DD}} - V_{\mathrm{T}n})}$$
$$R_p = \frac{1}{k'_p(\mathrm{W/L})_p(V_{\mathrm{DD}} - V_{\mathrm{T}p})} \tag{8.10}$$

These equations illustrate the general dependence that the drain–source resistance R is inversely proportional to the aspect ratio (W/L). However, the MOSFET is at best a nonlinear resistor, so that these are only rough estimates. It is important to note the MOSFET parasitic capacitances C_{GS}, C_{GD}, C_{SB}, and C_{DB} must be included in the switching models when performing transient analysis.

8.1.3 The Digital Inverter

An ideal digital inverter is shown in Figure 8.5. In terms of the Boolean variable A, the inverter accepts A and produces the complement $\overline{A}$. Electronic implementation of the inverter requires assigning voltage ranges for V_{in} and V_{out} to represent logic 0 and logic 1 states. These are chosen according to the DC

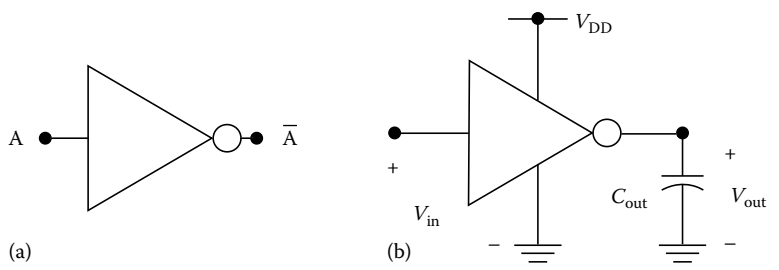


FIGURE 8.5 Basic inverter. (a) Ideal inverter symbol. (b) Electronic parameters.

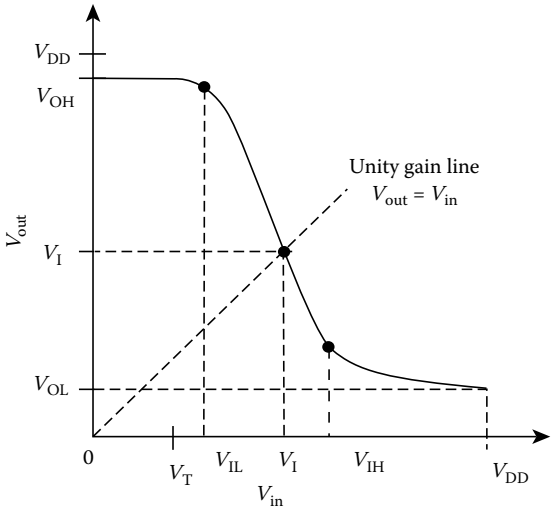


FIGURE 8.6 Inverter VTC.

voltage transfer characteristics (VTC) of the electronic circuit. A VTC is simply a plot of the output voltage V_{out} as a function of V_{in}; a general VTC is shown in Figure 8.6.

Consider first the output voltage V_{out}. The maximum value of V_{out} is denoted by V_{OH}, and is called the output high voltage. This is used to represent an ideal logic 1 output voltage. Conversely, the smallest value of V_{out} is denoted as V_{OL}, and is called the output low voltage. V_{OL} is the output logic 0 voltage. The logic swing of the inverter is then defined by $(V_{OH} - V_{OL})$. The range of input voltage V_{in} used to represent logic 0 and logic 1 input states is usually determined by points on the VTC at which the slope has a value of $(dV_{out}/dV_{in}) = -1$. Logic 0 voltages are those with values between 0 V and V_{IL}, the input low voltage. Similarly, voltages in the range from the input voltage V_{IH} to V_{OH} represent logic 1 input levels. The intersection of the VTC with the unity gain line defined by $V_{out} = V_{in}$ gives the inverter threshold voltage V_I; this represents the switching point of the circuit. The numerical value of V_{OH}, V_{OL} V_{IL}, V_{IHt}, and V_I are determined by the circuit and topology and the characteristics of the devices used in the circuit.

The transient characteristics of the gate are defined by two basic transition times. Figure 8.7 shows V_{in} and V_{out}, and the most important output switching intervals. The input voltage has been taken as an ideal step-like pulse. In a more realistic situation the input voltage is better approximated by a ramp or an exponential. The idealized pulse is used here because it allows a comparison among various circuits.

The most important switching properties of the inverter are the low-to-high time t_{LH}, and the high-to-low time t_{HL} as shown in Figure 8.7. These represent the minimum response times of the inverter.

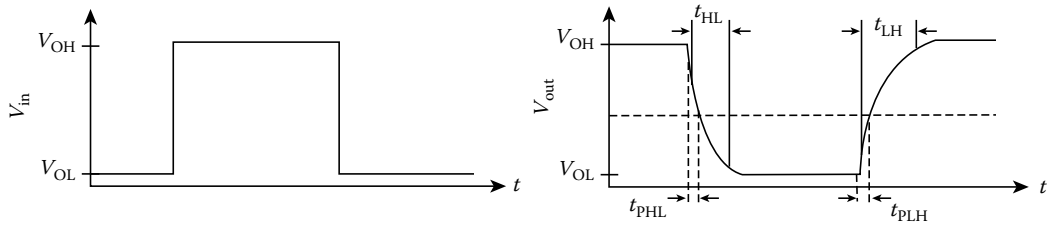


FIGURE 8.7 Inverter switching times.

Note that these time intervals are usually defined between the 10% and 90% voltages instead of the full logic swing. The maximum switching frequency is computed from

$$f_{\max} = \frac{1}{t_{\mathrm{LH}} + t_{\mathrm{HL}}} \tag{8.11}$$

The propagation delay t_{P} for the gate is the average time required for a change in the input to be seen at the output. It is computed using the time intervals t_{PHL} and t_{PLH} shown in the diagram from

$$t_{\mathrm{P}} = \left(\frac{1}{2}\right)(t_{\mathrm{PHL}} + t_{\mathrm{PLH}}) \tag{8.12}$$

Note that the transition times for this parameter are measured to the 50% voltage.

8.1.4 nMOS Logic Gates

Early generations of MOS logic circuits were based on a single type of MOSFET. The Intel 4004, for example, used only pMOS transistors, while subsequent microprocessor chips such as the Intel 8088, the Zilog Z80, and the Motorola 6800 used only *n*-channel MOSFETs. Although all current MOS-based designs are implemented in CMOS, which employs both nMOS and pMOS devices, it is worthwhile to examine nMOS-only logic circuits. This provides an introduction to the basic characteristics of MOS logic circuits, many of which are used in even the most advanced CMOS techniques.

Several types of inverter circuits can be constructed using *n*-channel MOSFETs. Three configurations are shown in Figure 8.8. Each circuit uses a switching transistor MD, known as the "driver," which is controlled by the input voltage V_{in}. The VTC is determined by the load device that connects the drain of MD to the power supply V_{DD}. The MD can be viewed as the switched "pull-down" device, while the load serves as the "pull-up" device. In Figure 8.8a, a simple linear resistor with a value R_{L} is used as the load. The circuits in Figure 8.8b use an enhancement-mode (defined with $V_{\mathrm{T}n} > 0$) nMOSFET biased into saturation, while Figure 8.8c has a depletion-mode nMOSFET (where $V_{\mathrm{T}n} < 0$) as an active load. Active loads provide better switching characteristics due to the nonlinearity of the device. In addition, MOSFETs are much smaller than resistors and process variations are not as critical because the circuit characteristics depend on the ratio of driver-to-load dimensions.

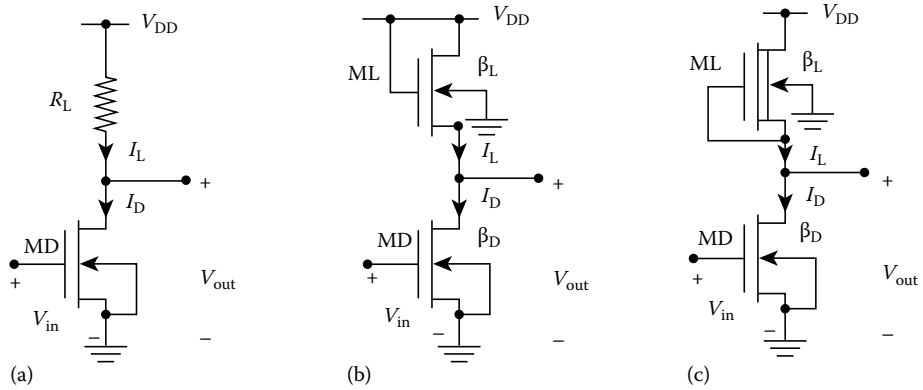


FIGURE 8.8 nMOS inverter circuits: (a) resistor load; (b) saturated enhancement model MOSFET load; and (c) depletion mode MOSFET load.

Although the three nMOS inverter circuits are similar in structure, they have distinct switching properties. Consider the output voltage swing. Figure 8.8a and c both have $V_{OH} \approx V_{DD}$, but the active load in Figure 8.8b gives

$$V_{OH} = V_{DD} - V_{TL} \tag{8.13}$$

with the threshold voltage computed from

$$V_{TL} = V_{T0n} + \gamma\left(\sqrt{2|\phi_F| + V_{OH}} - \sqrt{2|\phi_F|}\right) \tag{8.14}$$

This is referred to as a ''threshold voltage loss,'' and is due to the fact that the load must have a minimum gate–source voltage of $V_{GSL} = V_{TL}$ to be biased into the active mode. Obviously, $V_{OH} < V_{DD}$ for this circuit.

The value of $V_{OL} > 0$ is determined by a ratio of driver parameters to load parameters. In Figure 8.8a, this ratio is given by $R_L\beta_D$, which is inversely proportional to V_{OL}. This means a small V_{OL} requires that both the load resistance and the driver dimensions are large. In Figure 8.8b and c, V_{OL} is set by the driver-to-load ratio $\beta_R = (\beta_D/\beta_L) = (W/L)_D/(W/L)_L$; increasing β_R decreases V_{OL}. For the depletion MOSFET load circuit in Figure 8.8c, the design equation is given by

$$\beta_R = \frac{|V_{TL}|^2}{2(V_{DD} - V_{TD})V_{OL} - V_{OL}^2} \tag{8.15}$$

A condition of $\beta_R > 1$ is generally required to achieve a functional inverter, implying that the driver MOSFET is always larger than the load device. Also, note that is not possible to achieve $V_{OL} = 0$ because this requires an infinite driver-to-load ratio.

The transient switching characteristics are obtained by including the output capacitance C_{out} as shown in Figure 8.5 at the output. C_{out} consists of the input gate capacitance seen looking in the MOSFET of the next stage, and also has parasitic contributions from the MOSFETs and interconnects. By using a switch model for the driver, it is seen that the transient characteristics are determined by the time required to charge and discharge C_{out}. The high-to-low time t_{HL} represents the time it takes to discharge the capacitor through the driver MOSFET with a device transconductance value of β_D. A rough estimate is obtained using the RC time constant such that $t_{HL} \approx 2R_D\ C_{out}$, with R_D the equivalent resistance. Similarly, the low-to-high time t_{LH} is the time interval needed to charge C_{out} through the load device. With respect to Figure 8.8c has the best transient response such that $t_{LH} \approx 2R_L\ C_{out}$, where R_L represents the equivalent resistance of the load MOSFET. nMOS circuits in the mid-1980s had inverter transition times on the order of a few nanoseconds. Because the DC design requires that $\beta_R = (\beta_D/\beta_L) > 1$, and the drain–source resistance of a MOSFET is inversely proportional to β, these circuits exhibit nonsymmetrical switching times with $t_{LH} > t_{HL}$. The propagation delay times can be estimated using $t_{PHL} \approx R_D\ C_{out}$ and $t_{PLH} \approx R_L\ C_{out}$ because these are measured relative to the 50% voltage levels.

MOS-based logic design allows one to easily construct other logic functions using the inverter circuit as a guide. For example, adding another driver MOSFET in parallel gives the NOR operation, while adding a series connected driver yields the NAND operation; these are shown in Figure 8.9.

Complex logic gates for AOI (AND-OR-INVERT) and OAI (OR-AND-INVERT) canonical logic functions can be constructed using the simple rules

- nMOSFETs (or groups) in parallel provide the NOR operation
- nMOSFETs (or groups) in series provide the NAND operation

Examples are provided in Figure 8.10. It should be noted that this type of circuit structuring is possible because the drain and source are interchangeable. The main problem that arises in design complex

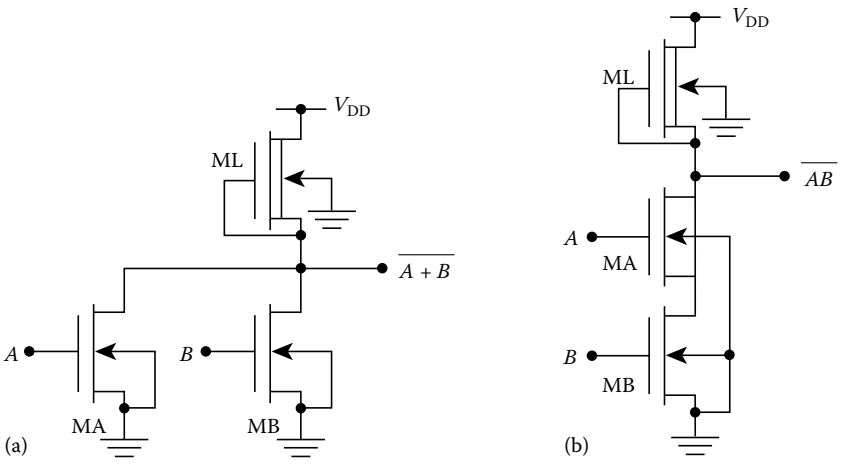


FIGURE 8.9 nMOS NOR and NAND gates: (a) two-input NOR gate and (b) two-input NAND gate.

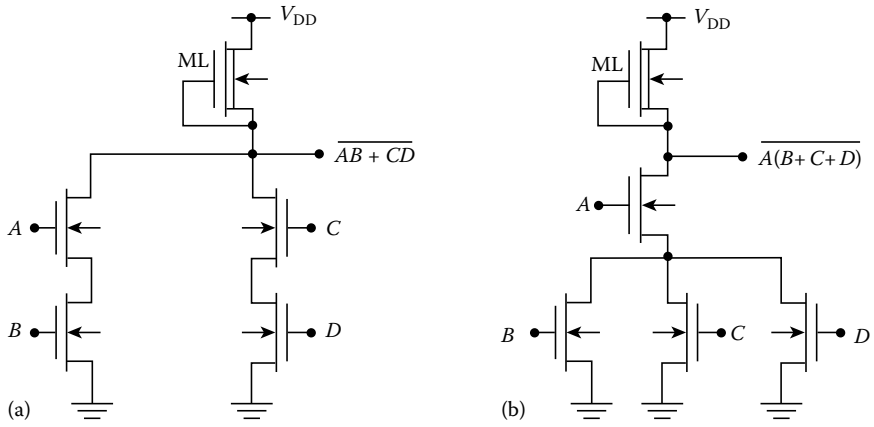


FIGURE 8.10 nMOS AOI logic gates.

nMOS logic gates is that the circuit requires large driver-to-load ratios to achieve small V_{OL} values. The switching FET arrays collectively act like a driver network that must be designed to have a large overall effective β value. Although parallel-connected MOSFETs are not a problem, the pull-down resistance of series-connected MOSFETs can be large unless the individual aspect ratios are increased. Satisfying this condition requires additional chip area, decreasing the logic density.

8.1.5 CMOS Inverter

A CMOS inverter is shown in Figure 8.11. This circuit uses a pair of transistors, one nMOS and one pMOS, connected with their gates together. When $V_{in} < V_{Tn}$, the pFET is active and the nFET is in cutoff. Conversely, when $V_{in} > (V_{DD} - V_{Tn})$, the nFET is active while the pFET is in cutoff. The two MOSFETs are said to form a complementary pair.

The complementary arrangement of the MOSFETs gives the circuit a full rail-to-rail output range, i.e., $V_{OL} = 0$ V and $V_{OH} = V_{DD}$. The devices are connected in such a way that terminal voltages satisfy

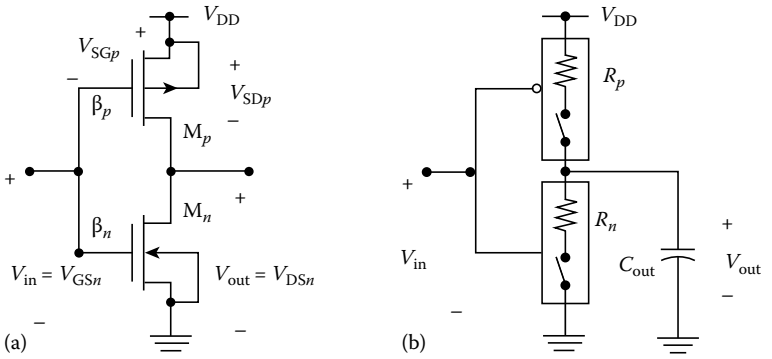


FIGURE 8.11 CMOS inverter: (a) circuit and (b) switch model.

$$\begin{aligned} V_{GSn} + V_{SGp} &= V_{DD} \\ V_{DSn} + V_{SDp} &= V_{DD} \end{aligned} \tag{8.16}$$

Note in particular the relationship between the gate–source voltages. Increasing the voltage on one transistor automatically decreases the voltage applied to the other. This provides the VTC with a very sharp transition, as shown in Figure 8.12. Moreover, the shape of the VTC is almost insensitive to the power supply value V_{DD}, which allows CMOS circuits based on this construction to be used with a range of values. The minimum value of V_{DD} is set by the device the device threshold voltages, and is usually estimated as being about 3 V_T. This is based on the input switching voltage V_{in}, and allows one V_T to switch the nFET, one V_T to switch the pFET, and one V_T for separation. Currently, V_T values equal ~0.5 V, so that the minimum V_{DD} is about 1.5 V. Because V_T is set in the fabrication, the minimum power supply used in low-voltage designs depends upon the process specifications. The maximum value of the power supply voltage is limited by the reverse breakdown voltages of the drain–bulk junctions. This is typically around 14–17 V.

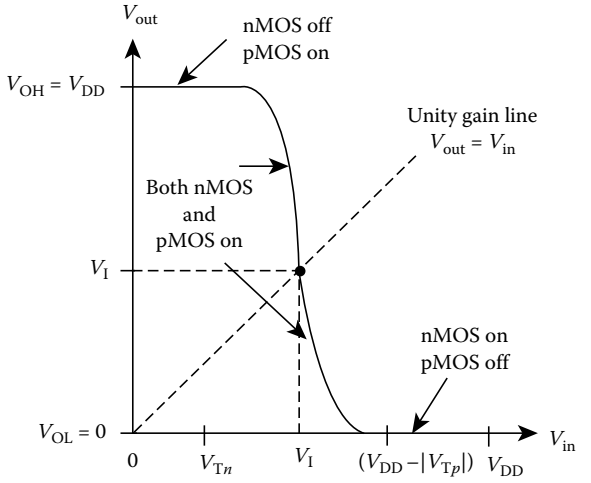


FIGURE 8.12 CMOS VTC.

Because the structure of the CMOS circuit automatically gives a full-rail output logic swing, the DC design of the gate centers around setting the inverter threshold voltage V_I. At this point, both FETs are saturated, and equating currents gives the expression

$$V_I = \frac{\sqrt{\beta_n/\beta_p}V_{Tn} + (V_{DD} - |V_{Tp}|)}{1 + \sqrt{\beta_n/\beta_p}} \tag{8.17}$$

This equation shows that V_I can be set by adjusting the ratio β_n/β_p. If $\beta_n = \beta_p$, and $V_{Tn} \approx |V_{Tp}|$, then $V_I \approx (V_{DD}/2)$. Increasing this ratio decrease the inverter switching voltage. If the nFET and pFET are of equal size, then $\beta_n > \beta_p$ (as $k'_n > k'_p$), and $V_I < (V_{DD}/2)$.

The transient characteristics are obtained by analyzing the charge and discharge current flow paths through the transistors. By using the switch model in Figure 8.11b, the primary time constants are

$$\begin{aligned} \tau_n &= R_n C_{out} = \frac{C_{out}}{\beta_n(V_{DD} - V_{Tn})} \\ \tau_p &= R_p C_{out} = \frac{C_{out}}{\beta_p(V_{DD} - |V_{Tp}|)} \end{aligned} \tag{8.18}$$

Analyzing the transitions with a step input voltage yields

$$\begin{aligned} t_{HL} &= \tau_n \left[\frac{2(V_{Tn} - V_0)}{(V_{DD} - V_{Tn})} + \ln\left(\frac{2(V_{DD} - V_{Tn})}{V_0} - 1 \right) \right] \\ t_{LH} &= \tau_p \left[\frac{2(|V_{Tp}| - V_0)}{(V_{DD} - |V_{Tp}|)} + \ln\left(\frac{2(V_{DD} - |V_{Tp}|)}{V_0} - 1 \right) \right] \end{aligned} \tag{8.19}$$

where $V_0 = 0.1\ V_{DD}$ is the 10% voltage. Noting once again that $k'_n > k'_p$, equal size transistors will give $t_{LH} > t_{HL}$. To obtain symmetrical switching, the pMOSFET must have an aspect ratio of $(W/L)_p = (k'_n > k'_p)\ (W/L)_n$. This illustrates that while the ratio of β-values sets the DC switching voltage V_I, the individual choices for β_n and β_p determine the transient switching times. In general, fast switching requires large transistors, illustrating the speed vs. area trade-off in CMOS design. The propagation delay time exhibits the same dependence.

Another interesting characteristic of the CMOS inverter is the power dissipation. Consider an inverter with stable logic 0 or logic 1 inputs. Because one MOSFET is in cutoff, the DC power supply current I_{DD} is very small, being restricted to leakage levels. The standby DC power dissipation is $P_{DC} = I_{DD}V_{DD} \approx 0$, so that static logic circuits do not dissipate much power under static conditions. Appreciable I_{DD} from the power supply to ground flows only during a transition. Dynamic power dissipation, on the other hand, occurs due to the charging and discharging of the output capacitance C_{out}. The dynamic power dissipation can be estimated by

$$P_{Dynamic} = C_{out}V_{DD}^2 f \tag{8.20}$$

where f is the switching frequency of the signal. Qualitatively, this is understood by noting that this is just twice the average stored energy multiplied by the frequency. This illustrates the important result that the power dissipation of a CMOS circuit increases with the switching frequency.

8.1.6 Static CMOS Logic Gates

Static logic gates are based on the inverter. The term ''static'' means that the output voltages (logic levels) are well defined as long as the inputs are stable. The nFET rules discussed for nMOS logic gates still apply

to CMOS. However, static logic gates provide an nFET and a pFET for every input. Proper operation requires that rules be developed for the pMOSFET array as follows:

- pMOSFETs (or groups) in parallel provide the NAND operation
- pMOSFETs (or groups) in series provide the NOR operation

When these rules are compared to the nMOS rules, it is seen that the nFET and pFET arrays are logical duals of one another (i.e., OR goes to AND, and vice versa).

An N-input static CMOS logic gate requires 2 N transistors. NAND and NOR gates are shown in Figure 8.13 using the rules; this type of logic is termed series-parallel, for obvious reasons. Examples of complex logic gates are shown in Figure 8.14. Note in particular the circuit in Figure 8.14b. This implements the XOR function by means of

$$A \oplus B = \overline{AB + \overline{A}\,\overline{B}} = A\overline{B} + \overline{A}B \tag{8.21}$$

Reductions of this type are often performed to work the AOI or OAI equation into a more familiar form.

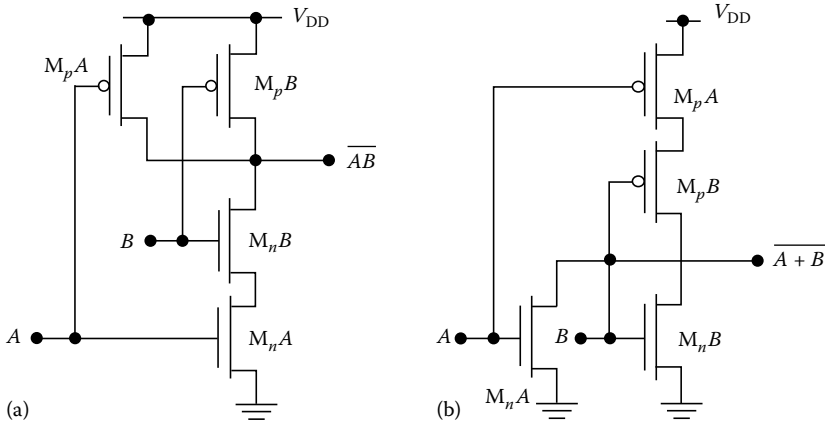


FIGURE 8.13 CMOS: (a) NAND and (b) NOR gates.

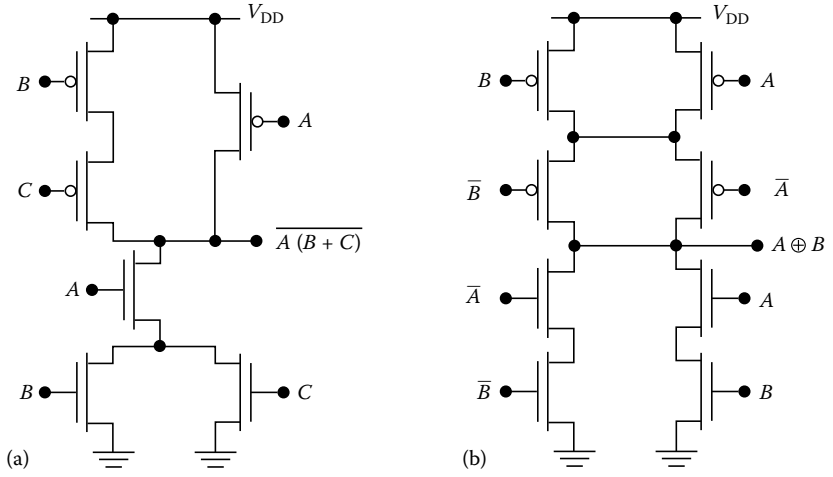


FIGURE 8.14 CMOS AOI logic examples: (a) AOI gate and (b) XOR circuit.

As seen from these examples, the logic function is determined by the placement of the nFETs and pFETs in their respective arrays. Electrically, the design problem centers around choosing the aspect ratios to achieve acceptable switching times. Because a MOSFET has a parasitic drain–source resistance that varies as $(1/\beta)$, series-connected transistor chains exhibit larger time constants than parallel-connected arrangements. Recalling that $R_n < R_p$ shows that for equal size devices, series chains of nFETs are preferable to the same number of series-connected pFETs. Consequently, NAND gates are used more frequently than NOR gates, and AOI logic functions with a small number of OR operations are better. It is also possible to expand to transistor arrays that are not of the series-parallel type, such as a delta configuration, but it is difficult to devise general design guidelines for these circuits.

Canonical CMOS static logic design is based on using pairs of nMOS and pMOS transistors. In modern very large scale integration (VLSI) design, the design complexity is limited by the interconnect (as opposed to the number of transistors), so that the need to connect every input to two transistors may result in problems in the chip layout. Pseudo-nMOS circuits provide an alternative to standard CMOS circuits. These logic gates implement logic using nFET arrays; however, the pMOS array is replaced by a single *p*-channel MOSFET that acts as a load device. Figure 8.15 shows an inverter and an AOI circuit implement based on pseudo-nMOS structuring. In both circuits, the load pMOSFET is biased active with $V_{SGp} = V_{DD}$ by grounding the gate. Although the circuits are simplified, two main problems arise with this type of circuit. First, the output low voltage V_{OL} is determined by the driver-to-load ratio $(\beta_n/\beta_p) > 1$, so that large driver nFETs are required. Second, if the input voltage is high, then the circuit dissipates DC power. Despite these drawbacks, pseudo-nMOS circuits may be useful in certain situations.

Transmission gates (TGs) provide another approach to implementing logic functions in CMOS. The properties of TGs are discussed in more detail in Section 8.2. However, because they are useful for certain types of logic gates, a short discussion has been included here. A basic TG consists of an nMOSFET and a pMOSFET in parallel, as shown in Figure 8.16a; the symbol in Figure 8.16b represents the composite structure. TGs act like voltage-controlled switches: logically, a condition of $C = 0$ gives an open switch, while $C = 1$ gives a closed switch. TGs can pass the full range of voltages (from 0 V to V_{DD}) in either direction; this is not possible with a single device, due to the threshold voltage characteristic discussed earlier in this section.

Figure 8.17 illustrates a simple 2:1 multiplexer (MUX) with two input lines, D_0 and D_1, and a control bit *S*. When $S = 0$, the upper TG is closed, and the output is $F = D_0$. Conversely, $S = 1$ closes the bottom TG, so $F = D_1$. The operation of this circuit is expressed by

$$F = \overline{S}D_0 + SD_1 \tag{8.22}$$

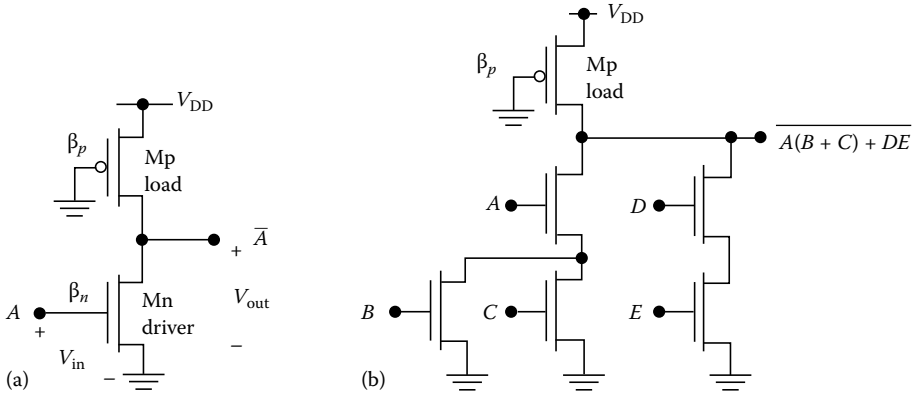


FIGURE 8.15 Pseudo-nMOS logic circuits: (a) Inverter and (b) AOI logic gate.

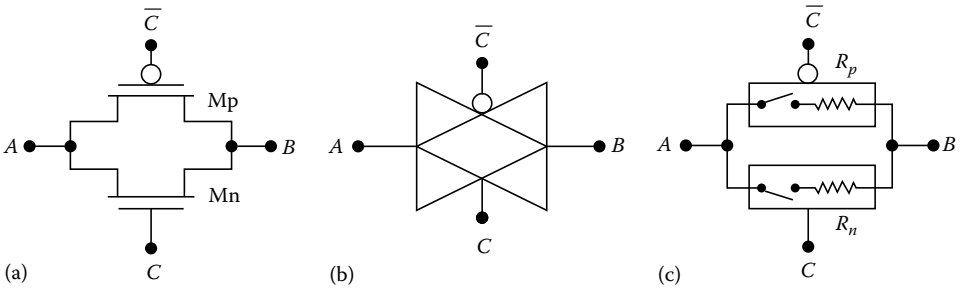


FIGURE 8.16 Transmission gate: (a) circuit; (b) symbol; and (c) switching model.

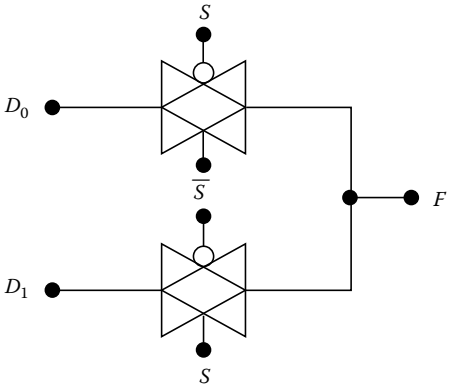


FIGURE 8.17 TG-based 2:1 multiplexer.

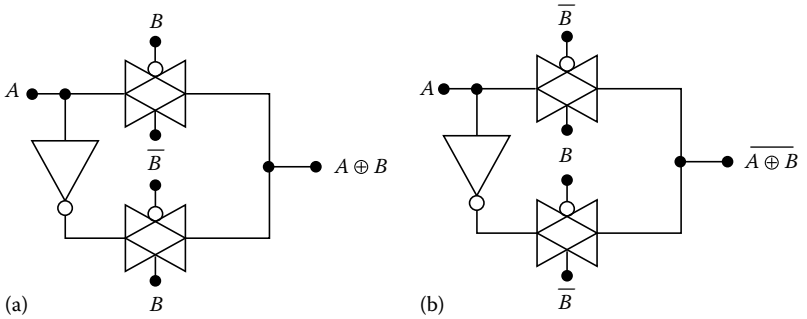


FIGURE 8.18 TG-based (a) XOR and (b) XNOR logic gates.

The circuit can be expanded easily to create larger multiplexers. For example, an 8:1 requires three select bits, and each of the eight lines will be a switching network using three TGs. Several other TG-based logic functions are popular in CMOS design. Figure 8.18 shows the exclusive-OR (XOR) and exclusive-NOR (XNOR) circuits. The primary drawbacks of TG-based logic circuits are that (1) the TG does not have a connection to the power supply, and acts as a parasitic RC element to the stage that drives it, and (2) the chip layout may become large, complicated, or both. In particular, (1) implies that TG circuits may be slower than equivalent functions designed using basic static CMOS techniques.

8.1.7 Dynamic CMOS Logic Gates

Dynamic CMOS logic gates are characterized as having outputs that are valid only for a limited time interval. Although this property inherently makes the circuit design more challenging, dynamic logic circuits can potentially achieve fast switching speeds. In general, dynamic circuits use parasitic capacitors in the MOS circuit to store charge Q. Because $Q = CV$, the presence or absence of charge corresponds to a logic 1 or logic 0 level, respectively, MOSFETs are used as voltage-controlled switches to "steer" the charge on and off the logic nodes. Several dynamic logic families have appeared in the literature, each having distinct characteristics. We now merely touch on some characteristics of a basic circuit which illustrates the important points.

Consider the dynamic logic circuit in Figure 8.19 for a three-input NAND gate. The transistors labeled MP and MN are controlled by the clock $\phi(t)$, and provide synchronization of the data flow. Note the presence of capacitors C_{out}, C_1, C_2, and C_3. These represent parasitic capacitances due to the transistors and interconnect, and are crucial to the operation.

The circuit is controlled by the timing provided by $\phi(t)$. When $\phi = 0$, MP is ON and MN is OFF. During this time, C_{out} is charged to a voltage $V_{out} = V_{DD}$, which is called a "precharge event." When ϕ changes to a level $\phi = 1$, MP is driven into cutoff, but MN is biased ON; the operation of the circuit during this time is termed a "conditional discharge event." If the inputs are set to $(A, B, C) = (1, 1, 1)$, then all three logic transistors, MA, MB, and MC, are ON, and C_{out} can discharge through these transistors and MN to a final voltage of $V_{out} = 0$ V. If at least one input is a logic 0, then C_{out} does not have a direct discharge path to ground. Ideally, V_{out} would stay at V_{DD}. However, charge leakage occurs across the reverse-biased drain–bulk *pn* junctions in the MOSFETs, eventually leading to a value of $V_{out} = 0$ V. Typically, the output voltage can be held only for a few milliseconds, thus leading to the name "dynamic circuit."

Another problem that arises in dynamic logic circuits is that of charge sharing. Consider the three-input NAND gate with inputs of $(A, B, C) = (0, X, X)$ during the precharge, where X is a do not care condition. The total charge transferred to the circuit from the power supply is

$$Q_T = C_{out} V_{DD} \tag{8.23}$$

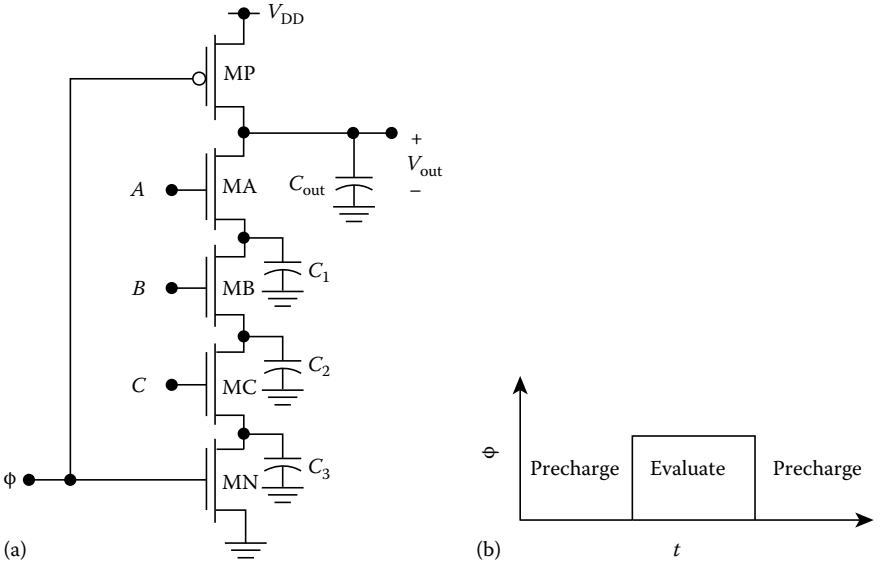


FIGURE 8.19 Dynamic CMOS logic gate: (a) three-input NAND gate and (b) timing intervals.

Now suppose that the inputs are switched to $(A, B, C) = (1, 1, 0)$ during the evaluation phase. MOSFETs MA and MB are ON, but MC is OFF, blocking the discharge path. Charge sharing occurs because the charge originally stored on C_{out} is now shared with C_1 and C_2. After the transients have decayed, the three capacitors are in parallel, Ignoring any threshold drop, they will share the same final voltage V_f such that

$$Q_T = (C_{out} + C_1 + C_2)V_f \tag{8.24}$$

Equating the two expressions for charge gives

$$V_f = \frac{C_{out}}{C_{out} + C_1 + C_2} V_{DD} < V_{DD} \tag{8.25}$$

To ensure that the output voltage remains at a logic 1 high voltage, the capacitors must satisfy the relation

$$C_{out} \gg C_1 + C_2 \tag{8.26}$$

The capacitance values are proportional to the sizes of the contributing regions, so that the performance is closely tied to the layout of the chip.

References

The material in this section is quite general. The references listed below are books in the field of digital MOS integrated circuits that provide further reading on the topics discussed here.

1. L. A. Glasser and D. W. Dobberpuhl, *The Design and Analysis of VLSI Circuits*, Reading, MA: Addison-Wesley, 1985.
2. H. Haznedar, *Digital Microelectronics*, Reading, MA: Addison-Wesley, 1991.
3. J. P. Uyemura, *Circuit Design for CMOS VLSI*, Norwell, MA: Kluwer Academic, 1992.
4. J. P. Uyemura, *Fundamentals of MOS Digital Integrated Circuits*, Reading MA: Addison-Wesley, 1988.

8.2 Transmission Gates

Robert C. Chang and Bing J. Sheu

A signal propagates through a transmission gate (TG) in a unique manner. In conventional logic gates, the input signal is applied to the gate terminal of an MOS transistor and the output signal is produced at the drain or the source terminal. In a TG, the input signal propagates between the source and the drain terminals through the transistor channel, while the gate voltage is held at a constant value. The TG is turned off if the voltage applied to the gate terminal is below the threshold voltage. The TG approach can be used in digital data processing to implement special switching functions with high performance as well as a small transistor count [1]. It also can be used in analog signal processing to act as a compact voltage-controlled resistor.

8.2.1 Digital Processing

8.2.1.1 Single Transistor Version

A TG can be constructed by a single nMOS or pMOS transistor, as shown in Figure 8.20. For an nMOS TG to pass a signal V_{in} to the output terminal, the selection signal S is set to the logic 1 value, i.e., the gate

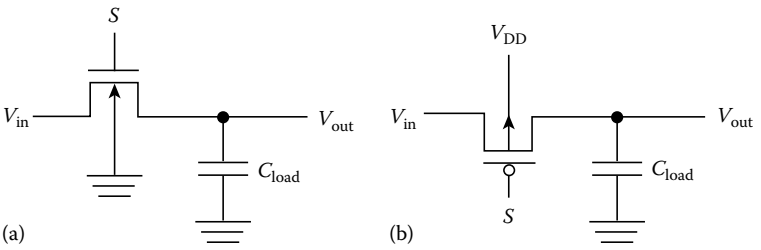


FIGURE 8.20 (a) nMOS TG and (b) pMOS TG.

voltage V_G is set to a high voltage value V_{DD}. If the input signal is also the V_{DD} value, the output voltage V_{out} is determined by [2],

$$V_{out}(t) = (V_{DD} - V_{thn})\left[\frac{t/\tau_{nc}}{1 + (t/\tau_{nc})}\right] \tag{8.27}$$

where

V_{thn} is the threshold voltage of the nMOS transistor with the body effect
τ_{nc} is the charging time constant which can be expressed as

$$\tau_{nc} = \frac{2C_{out}}{\mu_n C_{OX}(W/L)(V_{DD} - V_{thn})} \tag{8.28}$$

where

μ_n is the carrier mobility
C_{OX} is the per-unit-area capacitance value
W/L is the transistor aspect ratio

If time t goes to ∞, then V_{out} will approach $V_{DD} - V_{thn}$, which indicates that a threshold voltage loss occurs in the signal from the input node to the output node. This is due to the fact that V_{GS} must be greater than the threshold voltage to turn on the nMOS transistor. Owing to this voltage reduction, an nMOS TG can only transmit a ''weak'' logic 1 value. However, a logic 0 can be transmitted by an nMOS TG without penalty. In order to analyze this case, we set $V_{in} = 0$ and $V_{out}(t = 0) = V_{DD} - V_{thn}$. The output voltage V_{out} is determined by

$$V_{out}(t) = (V_{DD} - V_{thn})\left[\frac{2e^{-(t/\tau_{nd})}}{1 + e^{-(t/\tau_{nd})}}\right] \tag{8.29}$$

where the discharge time constant can be expressed as

$$\tau_{nd} = \frac{C_{out}}{\mu_n C_{OX}(W/L)(V_{DD} - V_{thn})} \tag{8.30}$$

Notice that V_{out} will approach zero as time goes to infinity. Input–output (I–O) characteristics of an nMOS TG are shown in Figure 8.21.

The schematic diagram of a pMOS TG is shown in Figure 8.20b. For a pMOS TG to pass a signal V_{in} to the output terminal, the selection signal S is set to the logic 0 value. To transmit a logic 0 value with the initial V_{out} value being V_{DD}, the expression for V_{out} is given as

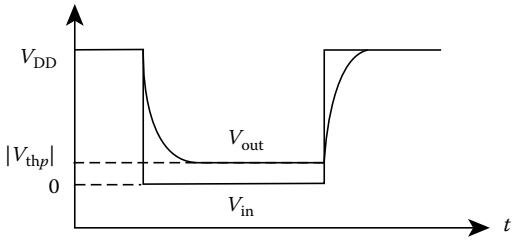


FIGURE 8.21 Characteristics of nMOS TG.

$$V_{out}(t) = |V_{thp}| + \frac{V_{DD} - |V_{thp}|}{1 + (V_{DD} - |V_{thp}|)(t/2\tau_{pd})} \tag{8.31}$$

where τ_{pd} is the discharging time constant for the pMOS TG. As time goes to infinity, V_{out} will approach $|V_{thp}|$, so that the pMOS TG can only transmit a "weak" logic 0 value. On the other hand, the pMOS TG can perfectly transmit a logic 1 value. To analyze this case, we set $V_{in} = V_{DD}$ and assume the initial V_{out} value as $|V_{thp}|$. The expression for V_{out} is given as

$$V_{out}(t) = V_{DD} - (V_{DD} - |V_{thp}|)\left(\frac{2e^{-(t/\tau_{pc})}}{1 + e^{-(t/\tau_{pc})}}\right) \tag{8.32}$$

where τ_{pc} is the charging time constant for the pMOS TG. The output voltage will approach V_{DD} as time goes to ∞. The transfer characteristics of the pMOS TG is shown in Figure 8.22.

8.2.1.2 Complementary Transistor Version

Figure 8.23 is the schematic diagram of a complementary transistor version of the TG which can be constructed by combining the characteristics of nMOS and pMOS TGs. The CMOS TG can transmit both the logic 0 and logic 1 values without any degradation. The voltage transmission properties of the single transistor and CMOS TGs are summarized in Table 8.1. The overall behavior of the CMOS TG can

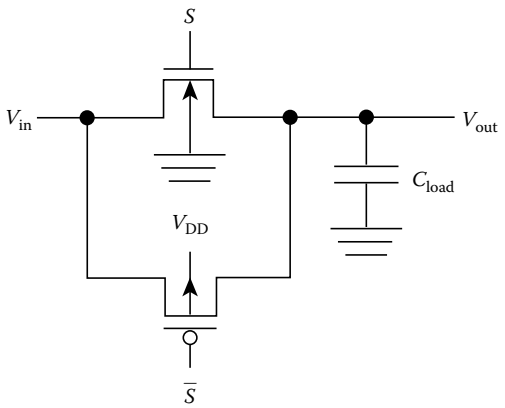


FIGURE 8.22 Characteristics of pMOS TG.

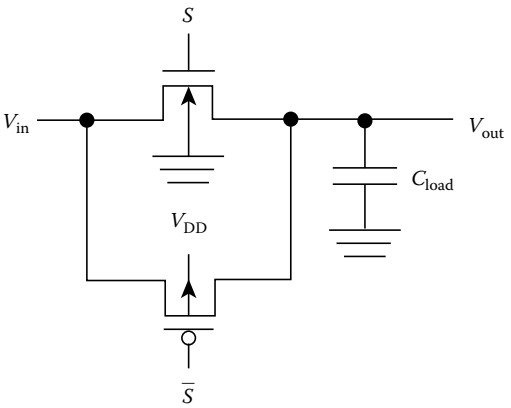


FIGURE 8.23 CMOS TG.

TABLE 8.1 Transmission Gate Characteristics

V_{out} V_{in} Type	$V_{in} = 0$ (Logic 0)	$V_{in} = V_{DD}$ (Logic 1)
nMOS	0	$V_{DD} - V_{thn}$
pMOS	$\|V_{thp}\|$	V_{DD}
CMOS	0	V_{DD}

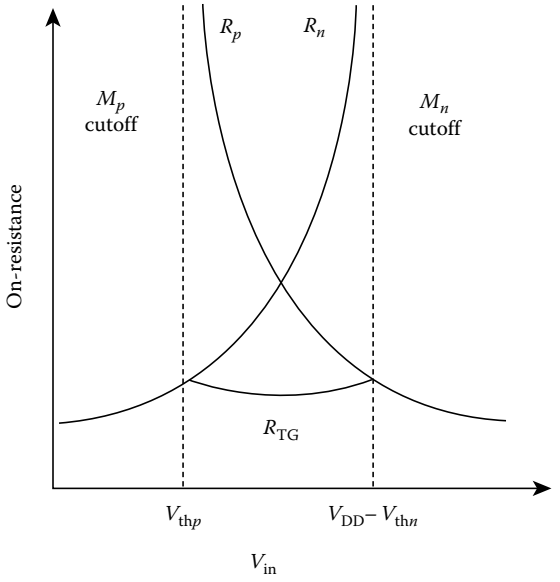


FIGURE 8.24 TG resistances.

be described as follows. When the selection signal S is low, both the nMOS and pMOS transistors are cut off. The output voltage V_{out} will remain at a high impedance state. When the selection signal S is high, both the nMOS and pMOS transistors are turned on and the output voltage will be equal to the input voltage.

Three regions of operation exist for a CMOS TG. In region 1 $V_{in} < |V_{thp}|$, then nMOS transistor is in the triode region and the pMOS transistor is in the cutoff region. Because the pMOS transistor is turned off, the total current, I_{tot}, is supplied by the nMOS transistor and I_{tot} decreases as V_{in} increases. In region 2 $|V_{thp}| < V_{in} < V_{DD} - V_{thn}$ both the nMOS and pMOS transistors are in the triode region. In this region, the nMOS transistor current decreases and the pMOS transistor current increases as V_{in} increases. Thus, I_{tot} is approximately a constant value. In region 3, $V_{in} > V_{DD} - V_{thn}$ the nMOS transistor is turned off and the pMOS transistor is in the triode region. The plot of the TG on-resistance is shown in Figure 8.24.

8.2.1.3 Pass-Transistor Logic

Pass-transistor logic is a family of logic which is composed of TG. Methods for deriving pass-transistor logic using nMOS TGs have been reported [3]. Figure 8.25 shows the schematic diagram of the pass-transistor logic in which a set of pass signals, P_is, are applied to the sources of the nMOS transistors and another set of control signals, C_is, are applied to the gates of the nMOS transistors.

The desired logic function F can be expressed as $F = C_1 \cdot P_1 + C_2 \cdot P_2 + \cdots + C_n \cdot P_n$. When C_is are high, P_is are transmitted to the output node. P_is can be logic 0, logic 1, true, or complement of the ith input variable X_i, or the high-impedance state Z. Constructing a Karnaugh map can help one to design the pass-transistor circuit. The pass function rather than the desired output values is put to the corresponding locations in the Karnaugh map. Then any variables that may act as a control variable or a pass variable are grouped.

For example, consider the design of a two-input XOR function. The truth table and the modified Karnaugh map of the XOR function are given in Tables 8.2 and 8.3, respectively. By grouping the A column when B is 0, and the $\overline{A}$ column when B is 1, the function can be expressed as

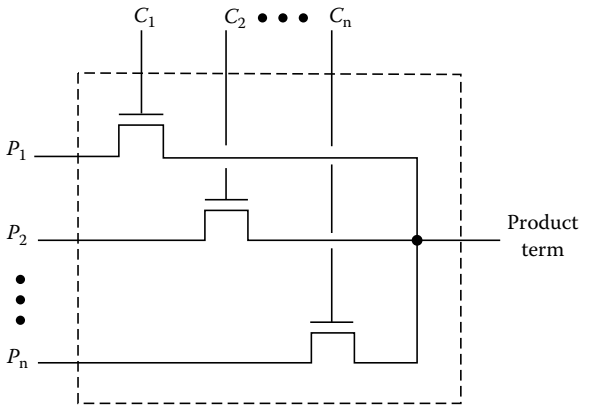


FIGURE 8.25 Model for pass transistor logic.

TABLE 8.2 Truth Table of XOR Function

A	B	A_B	Pass Function
0	0	0	$A+B$
0	1	1	$\overline{A}+B$
1	0	1	$A+\overline{B}$
1	1	0	$\overline{A}+\overline{B}$

TABLE 8.3 Modified Karnaugh Map for XOF Function

		B			
		0		1	
A	0	A		$\overline{A}$	
			B		B
	1	A		$\overline{A}$	
			$\overline{B}$		$\overline{B}$

$$F = \overline{B} \cdot A + B \cdot \overline{A} \tag{8.33}$$

where

B is a control variable

A is a pass variable

Figure 8.26a and b shows the schematic diagrams of nMOS and CMOS implementations of the XOR function. When the control variable B is with a logic 0 value, the pass variable A is transmitted to the output. When the control variable B is with a logic 1 value, the pass variable $\overline{A}$ is transmitted to the output. Another implementation of the XOR function is shown in Figure 8.26c.

It is not permitted to have groupings that transmit both true and false values of the input variable to the output simultaneously. The final expression must contain all the cells in the Karnaugh map. Note that

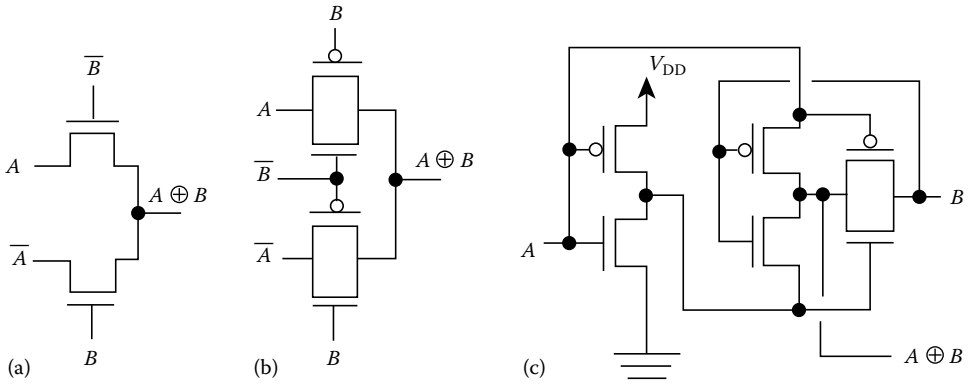


FIGURE 8.26 XOR gates: (a) nMOS version; (b) complementary version I; and (c) complementary version II.

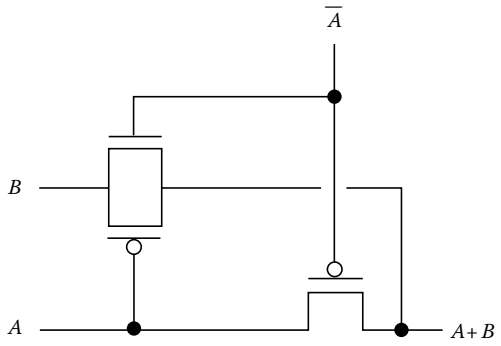


FIGURE 8.27 OR gates.

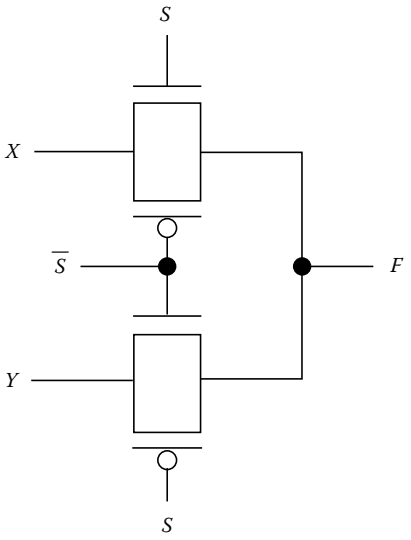


FIGURE 8.28 A two-input multiplexer.

the *p*-transistor circuit is the dual of the *n*-transistor circuit. Thus, the *p*-pass function must be constructed when a complementary version is required. In addition, the pass variable with logic 0 value is transmitted by the nMOS network in a complementary implementation while the pass variable with logic 1 value is transmitted by the pMOS network.

The OR function can be constructed by one pMOS transistor and one CMOS TG, as shown in Figure 8.27. When the input signal *A* is with the logic 0 value, the CMOS TG is turned on and the input signal *B* is passed to the output node. On the other hand, if the input signal *A* is with the logic 1 value, the pMOS TG is turned on and the logic 1 value of input signal *A* is transmitted to the output node. Because the pMOS TG can propagate a "strong" logic 1 value it is not necessary to use another CMOS TG.

TGs can be used to construct a multiplexer which selects and transmits one of the inputs to the output. Figure 8.28 is the circuit schematic diagram of a two-input multiplexer, which is composed of CMOS TGs. The output function of the two-input multiplexer is

$$F = X \cdot S + Y \cdot \overline{S} \tag{8.34}$$

If the selection signal *S* is at a logic 1 value, the input signal *X* is transmitted to the output. On the other hand, if the selection signal *S* is at a logic 0 value, the input signal *Y* is transmitted to the output. Multiplexers are important components in CMOS data manipulation structures and memory elements.

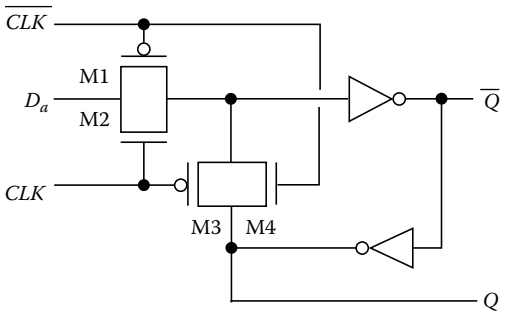


FIGURE 8.29 A CMOS *D* latch.

A basic *D* latch can be constructed by two TGs and two invertes, as shown in Figure 8.29. When the *CLK* signal is at a logic 0 value, pass transistors M1 and M2 are turned off so that the input signal D_a cannot be transmitted to the outputs Q and $\overline{Q}$. In addition, pass transistors M3 and M4 are turned on so that a feedback path around the inverter pair is established and the current state of Q is stored. When the *CLK* signal is at a logic 1 value, M1 and M2 are turned on and M3 and M4 are turned off. Thus, the output signal Q is set to the input signal D_a and $\overline{Q}$ is set to $\overline{D}_a$. Because the output signal Q will follow the change of input signal D_a when the *CLK* signal is high, this circuit is a positive level-sensitive *D* latch. A positive edge-trigger register or so-called *D* flip-flop can be designed by combining one positive level-sensitive *D* latch and one negative level-sensitive *D* latch. By cascading *D* flip-flops, a shift register can be constructed.

TGs can be used in the design of memory circuits. A typical random access memory (RAM) architecture consists of one row/word decoder, one column/bit decoder, and memory cells. The memory cells used in RAMs can be categorized into static cells and dynamic cells. Memory data/charges are stored on the latches in static cells, while on the capacitors in dynamic cells. The static random access memories (SRAMs) are not forced to include the refresh circuitry and are faster than the dynamic random access memories (DRAMs). However, the size of SRAM cells is much larger than that of DRAM cells. The most commonly used circuit in the design of SRAM cells is the six-transistor circuit shown in Figure 8.30. Four transistors are used to form two cross-coupled inverters. The other two transistors, M1 and M2, are TGs to control the read/write operation of the memory cell. If the word line is not selected, the data stored on the latch will not change as long as the leakage current is small. If the word line is selected, the transistors M1 and M2 are turned on. Through the *bit* and $\overline{bit}$ lines, data can be written into the latch or the stored data can be read out by the sense amplifier. TGs can also be found in the four-transistor DRAM cell circuit, as shown in Figure 8.30b. When the *Read* line is selected, pass transistor M1 is turned on and the data stored on the capacitor C_1 are read out. When the *Write* line is selected, pass transistor M2 is turned on and the data from *data_W* line are written into the cell.

Figure 8.31 is the circuit schematic diagram of a TG adder, which consists of four TGs, four inverters, and two XOR gates [4]. The *SUM* output, which represents $A \oplus B \oplus C$, is constructed by a multiplexer controlled by $A \oplus B$ and its complement. Notice that when $A \oplus B$ is false, the *CARRY* output equals *A* or *B*. Otherwise, *CARRY* output takes the value of input signal *C*. Although the TG adder has the same

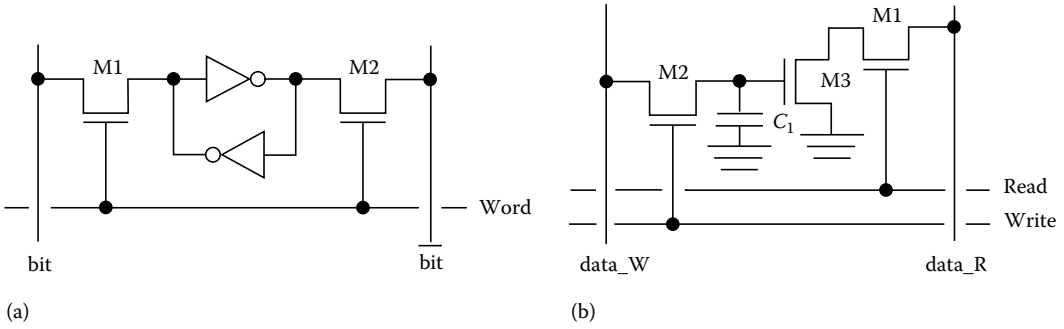


FIGURE 8.30 (a) SRAM cell and (b) DRAM cell.

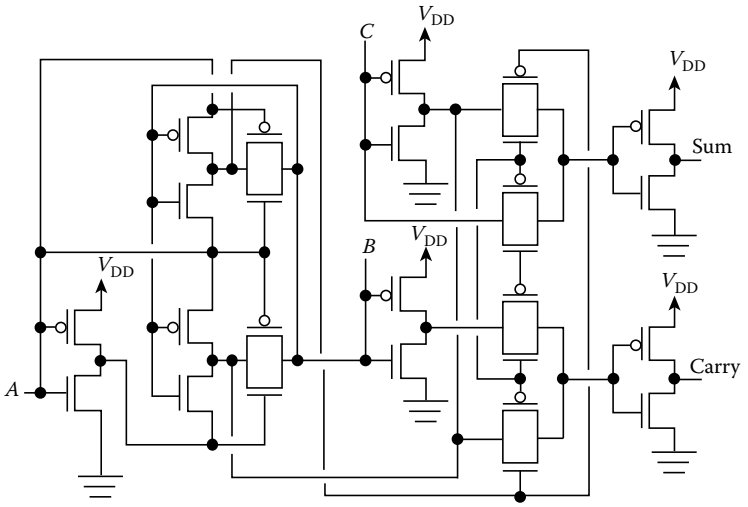


FIGURE 8.31 TG adder.

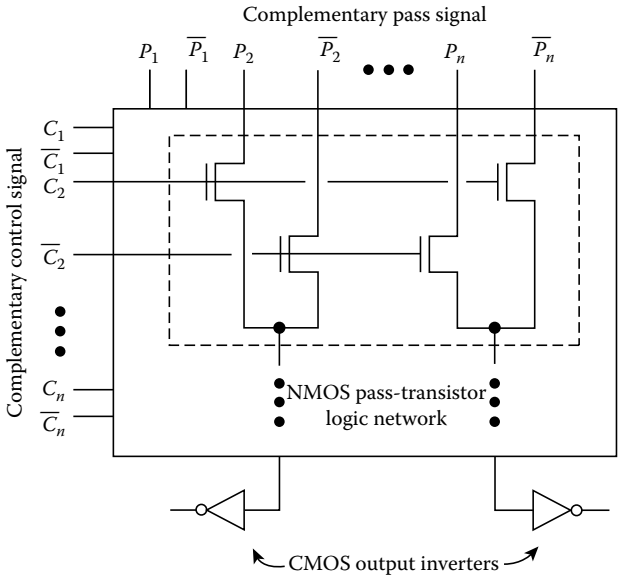


FIGURE 8.32 Schematic structure of the basic CPL circuit.

number of transistors as the combinational adder, it has the advantage of having noninverted *SUM* and *CARRY* output signals and an equal delay time for the *SUM* and *CARRY* output signals.

Another form of differential CMOS logic, complementary pass-transistor logic (CPL), has been developed and utilized on the critical path to achieve very high speed operation [5]. Figure 8.32 is the circuit schematic diagram of the basic CPL structure using an nMOS pass-transistor logic organization. The CPL is constructed by an nMOS pass-transistor logic network, complementary inputs and outputs, and CMOS output inverters. As the nMOS pass transistor will transmit a logic 1 signal with one threshold voltage reduction, the output signals must be amplified by the CMOS inverters which can shift the logic threshold voltage and drive a large capacitive load. One attractive feature of the CPL design

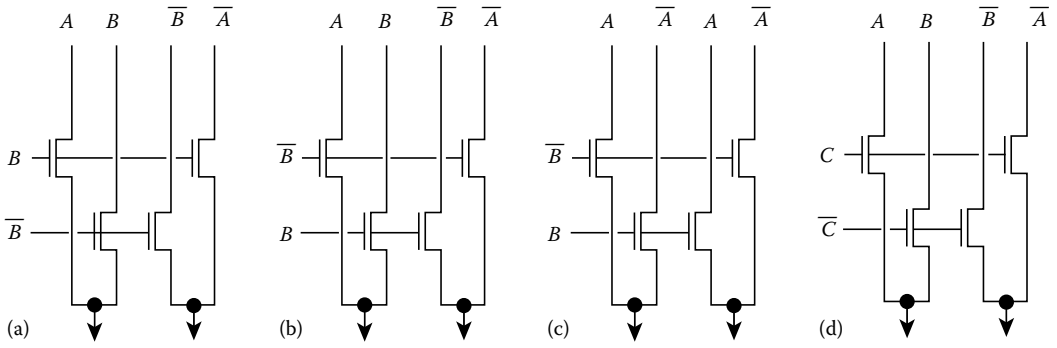


FIGURE 8.33 CPL circuit modules: (a) AND/NAND; (b) OR/NOR; (c) XOR/XNOR; and (d) Wire-AND/NAND.

is that complementary outputs are generated by the simple four-transistor circuits. Because inverters are not required in CPL circuits, the number of critical-path gate stages can be reduced.

Figure 8.33 shows the schematic diagrams of four basic CPL circuit modules: an AND/NAND module, an OR/NOR module, an XOR/XNOR module, and a wired-AND/NAND module [5]. By combining these four circuit modules, arbitrary Boolean functions can be constructed. These modules have an identical circuit schematic and are distinguished by different arrangements of input signals. This property of CPL is quite suitable for master-slice design.

The schematic diagram of a CPL full adder is shown in Figure 8.34. Both the circuitry to produce the *SUM* output signal and the circuitry to produce the *CARRY* output signal are constructed from basic CPL modules. The *SUM* circuitry consists of two XOR/XNOR modules, while the *CARRY* circuitry consists of three wired-AND/NAND modules. The CMOS output inverters are fixed ''overhead'' because they are required whether the circuit has one, two, or many inputs. Thus, designing with a complex Boolean function in a CPL gate is preferred to minimize the delay time and overall device count.

Figure 8.35 is the block diagram of a 16×16 bit multiplier, which is constructed by using a parallel multiplication architecture. A carry-look-ahead (CLA) adder and a Wallace-tree adder array are used to minimize the critical-path gate stages. The number of transistors in the CPL multiplier is less than that in a full CMOS counterpart [6].

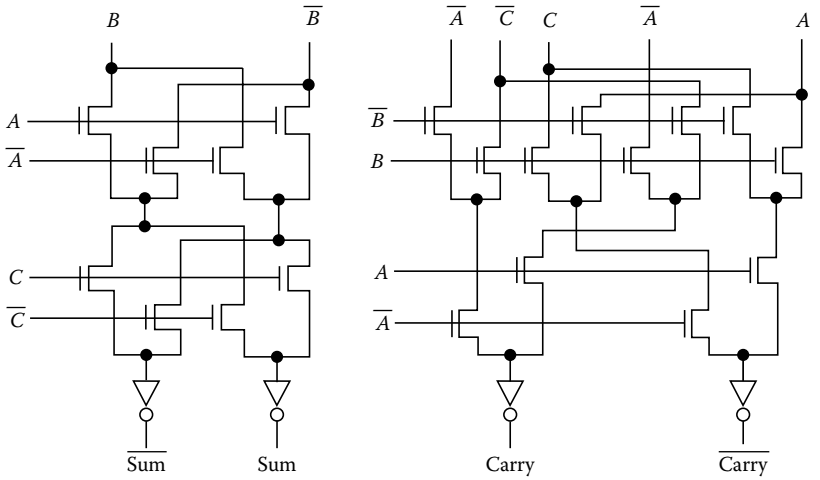


FIGURE 8.34 CPL full adder circuit.

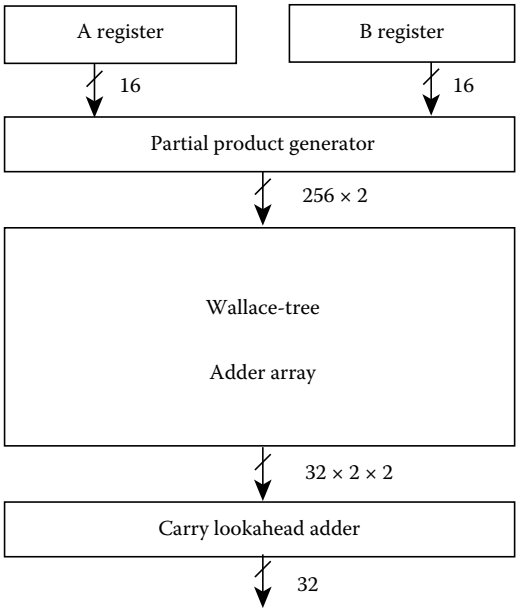


FIGURE 8.35 Block diagram of the 16 × 16 bit multiplier.

Due to the continued device miniaturization and the recent drive of portable systems, VLSI systems have been pushing toward low-voltage, low-power operation. Various techniques from system level to device level were developed to reduce the operating voltage and the power consumption of the VLSI circuits [7,8]. The low-power design can be addressed at four levels: algorithm, architecture, logic style, and integration. At the logic design level, capacitive loads are to be reduced and the number of charging/discharging operations are to be minimized. A CPL is one of the most attractive logic families that can achieve very low power consumption. The input capacitance in CPL is about half that of the CMOS configuration because pMOS can be eliminated in logic organization. Therefore, CPL can achieve a higher speed and dissipate less power. Experimental results [5] show that for the same delay time of the CMOS full adder operating at 5 V, the CPL adder requires only a 2 V supply. As the supply voltage decreases, the delay time will increase, but the power-delay product will decrease. Hence, it is desirable to operate at the slowest allowable speed to reduce power dissipation. Experimental results indicate that performance of CPL logic style is better than the conventional CMOS logic style from the viewpoint of power consumption.

8.2.2 Analog Processing

8.2.2.1 MOS Operational Amplifier Compensation

The frequency stabilization of a basic two-stage CMOS amplifier can be achieved by using a pole-splitting capacitor C_C [9]. The pole p_1 due to the capacitive loading of the first stage is pushed down to a very low frequency, and the pole p_2 due to the capacitance at the output node of the second stage is pushed to a very high frequency. However, a right-half-plane zero is introduced by the feedthrough effect of the compensation capacitor C_C. It will degrade the stability of the op-amp and make the second stage behavior like a short-circuited configuration at high frequencies. In order to remove the effects of the zero, a source follower can be inserted in the path from the output back through the compensation capacitor. Another approach is to insert a nulling resistance, R_Z, in series with the compensation capacitor. If R_Z is set to $1/g_{M2}$, where g_{M2} is the transconductance of the second stage, the zero

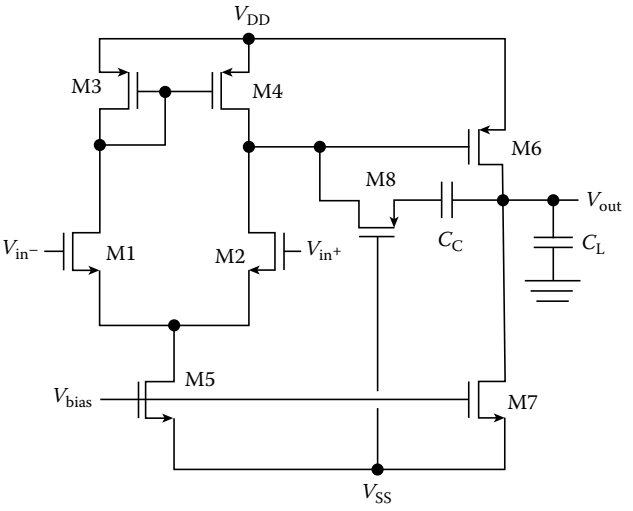


FIGURE 8.36 CMOS op-amp.

vanishes and the feedthrough effect is cancelled out. A single transistor or complementary version of the TG can be used to implement R_Z. Figure 8.36 is the schematic diagram of a basic two-stage op-amp supplemented by a feedback branch (M8, C_C) for compensation [10]. Capacitance C_L is the load capacitance to be driven by the amplifier. The pMOS TG M8 is biased in the triode region and provides the equivalent resistance.

TGs can be used to construct the cascode configuration of an op-amp. A fully differential folded-cascode op-amp is shown in Figure 8.37 [10]. The output cascode stage consists of TGs M5 to M10.

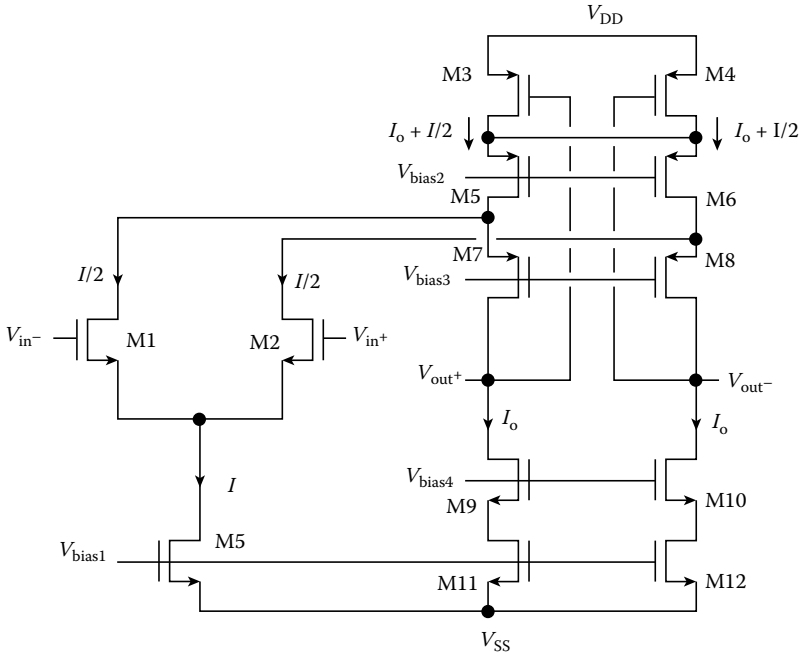


FIGURE 8.37 A fully differential CMOS op-amp with stabilized DC output level.

A high output impedance can be achieved by using the split-load arrangement. The bias voltage V_{bias1} establishes the bias current I of the input stage, and a bias current I_{o} in the output transistors M7 to M12. Thus, each transistor of M3 to M6 has a stabilized current $I_{\text{o}} + I/2$. The source voltages of M5 and M6 are stabilized because they conduct stabilized currents and their gate voltages are fixed at V_{bias2}. This fixes $|v_{\text{DS}\,3}|$ and $|v_{\text{DS}\,4}|$. Let transistors M3 and M4 have the same W/L ratio and bias them in the triode region by choosing a suitable value for V_{bias2}. If the output common-mode voltage $v_{\text{o,c}}$ drops, the resistance of M3 and M4 reduces, which increases $|v_{\text{GS}\,5}|$ and $|v_{\text{GS}\,6}|$. Because the current in M5 and M6 remains unchanged, $|v_{\text{DS}\,5}|$ and $|v_{\text{DS}\,6}|$ are forced to decrease. Then, the drain voltages of M5 and M6 increase, which increases $|v_{\text{GS}\,7}|$ and $|v_{\text{GS}\,8}|$. Therefore, $|v_{\text{DS}\,7}|$ and $|v_{\text{DS}\,8}|$ reduce which forces v_{o}^{+} and v_{o}^{-} to rise. The common-mode voltage $v_{\text{o,c}}$ is thus increased. This approach can increase the common-mode rejection ratio (CMRR) of the op-amp. The negative feedback scheme tends to keep $v_{\text{o,c}}$ at a constant value. It means that the small-signal common-mode output is zero or a very small value. Thus, a high CMRR is achieved.

8.2.2.2 Transimpedance Compensation

The optical receiver is an important component of the optical fiber communication system. One of the basic modules in a high performance optical receiver is the low-noise preamplifier. Several approaches are available to design the preamplifier. One approach is to use the transimpedance design which can avoid the equalization and limited dynamic range problems by using negative feedback. TGs can be used to provide the feedback resistance for a transimpedance amplifier.

A complete preamplifier circuit schematic is given in Figure 8.38 [11]. This circuit consists of three gain stages and two TGs. Each gain stage is composed of a pMOS current source achieving a common-source amplification with a folded nMOS load. One TG transistor, M10, functions as a feedback resistor and the other, M11, functions to implement the automatic gain control function. The gate voltage of M10 is derived from another circuit which minimizes the temperature and power supply dependence of the feedback resistance [11]. Transistor M11 is controlled by the automatic gain control voltage [12] and is normally off. If the input current to the preamplifier forces the output voltage out of its linear range, M11

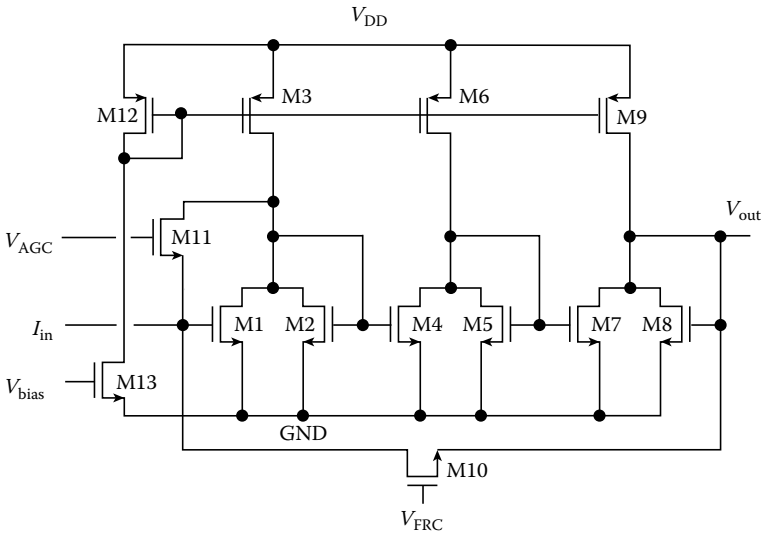


FIGURE 8.38 Circuit schematic of a preamplifier.

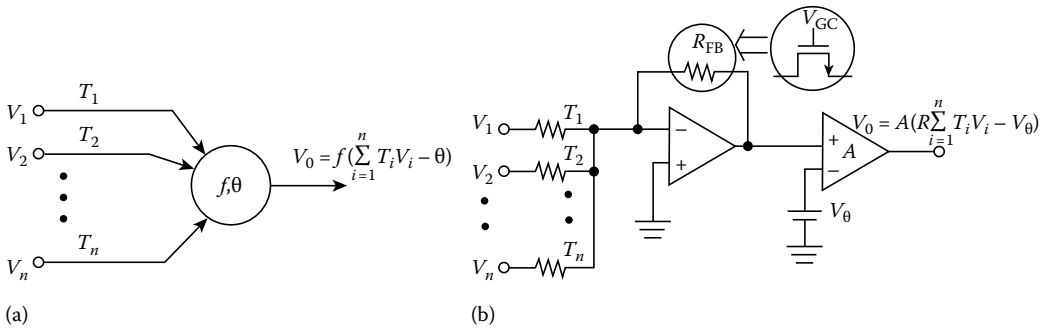


FIGURE 8.39 Neuron and synapse operation: (a) mathematical model and (b) analog circuit model with adjustable gain.

is turned on and begins to shunt current away from the feedback resistor and into the first stage output node.

With recent progress in intelligent information processing, artificial neural networks can be used to perform several complex functions in scientific and engineering applications, including classification, pattern recognition, noise removal, optimization, and adaptive control [13]. Design and implementation of VLSI neural networks have become a very important engineering task. The basic structure of an artificial neural network consists of a matrix of synapse cells interconnecting an array of input neurons with an array of output neurons. The inputs, V_i, are multiplied by weight values, T_i, of the synapses. The results of the multiplication are summed and compared to the threshold value θ_i in the output neurons. Schematic diagrams of a mathematical model of a neuron and its electronic counterpart are shown in Figure 8.39. The circuit in Figure 8.39b uses a gain-controllable amplifier in which the voltage gain is controlled by changing the feedback resistance. The feedback resistor R_{FB} can be constructed by the TG structure so that feedback resistance can be adjusted by the gain-control voltage V_{GC} [14].

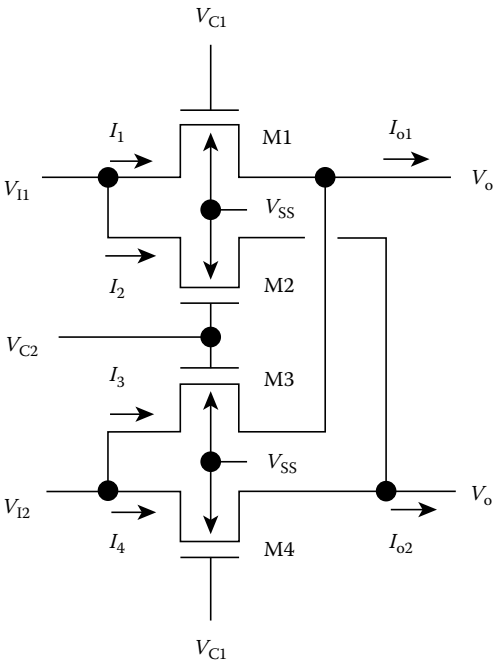


FIGURE 8.40 Double-MOS implementation of a differential AC resistor.

8.2.2.3 Continuous-Time Filters

Resistors are important components in the construction of continuous time filters [15]. However, the implementation of resistors by integrator circuit (IC) fabrication technologies was found to be lacking in several areas of performance. The TG can be used to realize active resistance. For example, a double-MOS differential configuration, shown in Figure 8.40, is used to implement a differential AC resistor [16].

This circuit consists of four nMOS TGs. Not only can it linearize the AC resistor, but it can also eliminate the effects of the bulk-source voltage [17]. To determine the AC resistance, assume that all the transistors are matched and are biased in the triode region.

The current I_{o1} and I_{o2} can be expressed as

$$\begin{aligned} I_{o1} &= I_1 + I_3 \\ &= \mu_n C_{OX}(W/L)\left[(V_{C1} - V_0 - V_{thn})(V_{I1} - V_0) - (1/2)(V_{I1} - V_0)^2\right] \\ &\quad + \mu_n C_{OX}(W/L)\left[(V_{C2} - V_0 - V_{thn})(V_{I2} - V_0) - (1/2)(V_{I2} - V_0)^2\right] \end{aligned} \tag{8.35}$$

$$\begin{aligned} I_{o2} &= I_2 + I_4 \\ &= \mu_n C_{OX}(W/L)\left[(V_{C2} - V_0 - V_{thn})(V_{I1} - V_0) - (1/2)(V_{I1} - V_0)^2\right] \\ &\quad + \mu_n C_{OX}(W/L)\left[(V_{C1} - V_0 - V_{thn})(V_{I2} - V_0) - (1/2)(V_{I2} - V_0)^2\right] \end{aligned} \tag{8.36}$$

Equations 8.35 and 8.36 can be combined to determine the differential current

$$I_{o1} - I_{o2} = \mu_n C_{OX}(W/L)[(V_{C1} - V_{C2})(V_{I1} - V_{I2})] \tag{8.37}$$

Thus, r_{ac} is given by

$$r_{ac} = \frac{V_{I1} - V_{I2}}{I_{o1} - I_{o2}} = \frac{1}{\mu_n C_{OX}(W/L)(V_{C1} - V_{C2})} \tag{8.38}$$

Because all transistors are required to be biased in the triode region, Equation 8.38 holds when

$$V_{I1}, V_{I2} \leq \min\left[V_{C1} - V_{thn}, V_{C2} - V_{thn}\right] \tag{8.39}$$

The double-MOSFET differential resistor is really a transresistance, thus, it can be applied only to differential-in, differential-out op-amps.

8.2.2.4 Switched-Capacitor Circuits

Switched-capacitor circuits make use of TGs in processing the analog signals [10,16]. This approach uses switches and capacitors and is in discrete time. If the clock rate is much higher than the signal frequency, an AC resistor can be implemented by combining switches and capacitors. The equivalent resistance is dependent only on the clock rate and the capacitor. The circuit schematic diagram of the direct digital integrator (DDI) is shown in Figure 8.41. The resistance is realized by two MOS switches and one capacitor. The difference equation can be expressed as

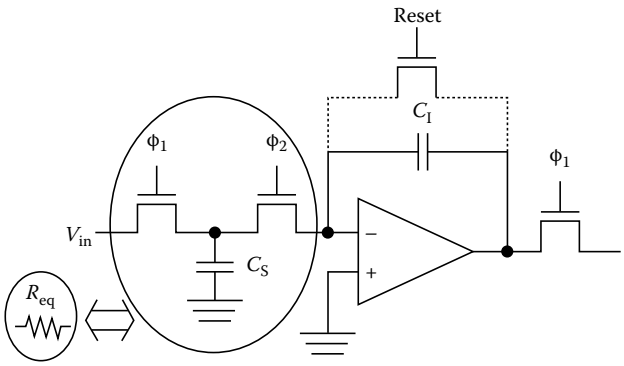


FIGURE 8.41 Direct digital integrator.

$$v_{0,n+1} = v_{0,n} - \frac{C_S}{C_I} v_{in} \tag{8.40}$$

After taking the z-transform, the new expression becomes

$$z \cdot V_0(z) = V_0(z) - \frac{C_S}{C_I} V_{in}(z) \tag{8.41}$$

By rearranging the various terms, the transfer function of the DDI can be expressed as

$$\frac{V_0(z)}{V_{in}(z)} = -\frac{C_S}{C_I} \cdot \frac{z^{-1}}{1 - z^{-1}} \tag{8.42}$$

By setting $z = e^{j\omega T}$, the frequency response can be determined:

$$\frac{V_0}{V_{in}}(j\omega) = -\frac{C_S}{C_I} \frac{1}{j\omega T} \cdot \frac{\omega T/2}{\sin(\omega T/2)} \cdot e^{-j\omega T/2} \tag{8.43}$$

where T is the period of the clock. In Equation 8.43, the first term corresponds to an ideal integrator, the second term contributes to the magnitude error, and the third term is the phase error. Because the frequency response of the ideal integrator is $-[j\omega R_{eq} C_I]^{-1}$, the equivalent resistance value is determined by

$$R_{eq} = \frac{T}{C_S} \tag{8.44}$$

A ladder network can be constructed by cascading the DDI integrators. In the ladder network all cascaded stages sample the input signal at clock Φ_1 and transform the signal at clock Φ_2, where Φ_1 and Φ_2 are nonoverlapping clock signals. This clocking scheme induces the extra half-cycle phase delay. This phase error can cause extra peaking in frequency response and generate cyclic response. In order to remove the excess phase, other integrators, such as lossless digital integrators (LDI) or bilinear integrators, can be used. In an LDI ladder network, the odd-number stages sample the input signal at clock Φ_1 and transform the signal at clock Φ_2, while the even-number stages sample the input signal at clock Φ_2 and transform the signal at clock Φ_1. Thus, the frequency response of an LDI integrator can be expressed by

$$\frac{V_0}{V_{in}}(j\omega) = -\frac{C_S}{C_I} \frac{1}{j\omega T} \cdot \frac{\omega T/2}{\sin(\omega T/2)} \tag{8.45}$$

Figure 8.42 is the circuit schematic diagram of the bottom-plate differential-input LDI. Output of an LDI integrator is more insensitive to parasitic components.

Figure 8.43 is the circuit schematic diagram of a differential bilinear integrator. The transfer function of the bilinear integrator is

$$\frac{V_0^+ - V_0^-}{V_{in}^+ - V_{in}^-} = \frac{C_S}{C_I} \cdot \frac{1 + z^{-1}}{1 - z^{-1}} \tag{8.46}$$

As the output of the bilinear integrator does not change during clock Φ_1, it can be used to feed another identical integrator.

TGs can be used to initialize the switched-capacitor circuits. For example, capacitor C_I in Figure 8.41 is to perform the integration function and to be reset or discharged before operation. An nMOS TG can be

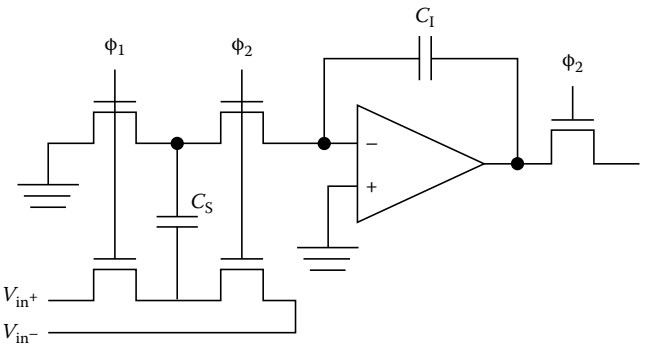


FIGURE 8.42 Bottom-plate differential-input lossless digital integrator.

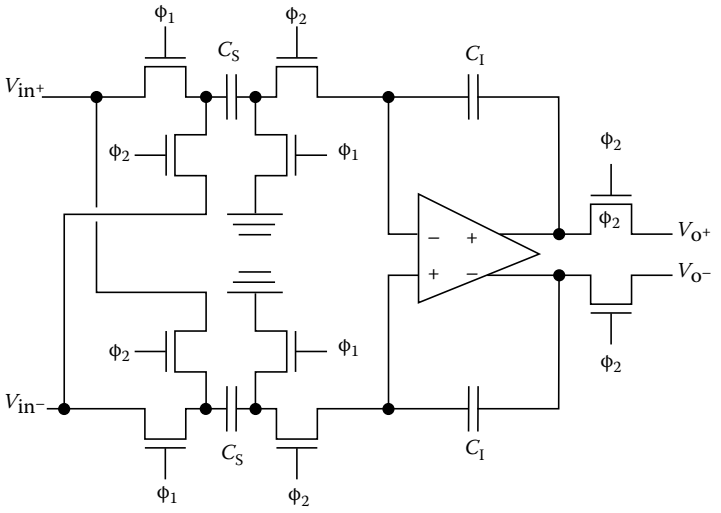


FIGURE 8.43 Differential bilinear integrator.

put in parallel with the capacitor C_I. Before normal operation, the TG is turned on and the capacitor C_I is discharged so that the initial capacitor voltage value is reset to zero.

The accuracy of switched-capacitor circuits is disturbed by charge injection when the controlling switch turns off. The turn-off of an MOS switch consists of two phases. The gate voltage is higher than the transistor threshold voltage V_{th} during the first phase. A conduction channel extends from the source to the drain of the transistor. As the gate voltage decreases, mobile carriers exit through both the drain and the source terminals and the channel conduction decreases. During the second phase, the gate voltage is smaller than V_{th} and the conduction channel no longer exists. The coupling between the gate and the data-holding node is only through the gate-to-diffusion overlap capacitance. The following analysis is focused on the switch charge injection due to the first phase of the switch turn-off.

Figure 8.44 is the circuit schematic corresponding to the general case of switch charge injection. Capacitance C_L is the lumped capacitance at the data-holding node. Capacitance C_S could be the lumped capacitance associated with the amplifier output node, while resistance R_S could be the output resistance of an op-amp. Let C_G represent the total gate capacitance of the switch, including both the channel

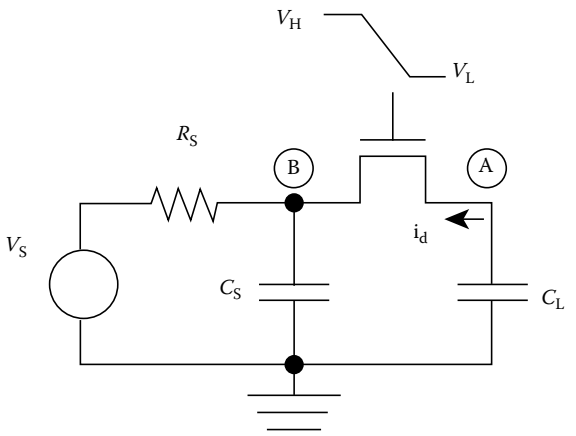


FIGURE 8.44 Circuit for analysis of switch charge injection.

capacitance and gate-to-drain/gate-to-source overlap capacitances. Kirchhoff's current law at node *A* and node *B* requires

$$C_L \frac{dv_L}{dt} = -i_d + \frac{C_G}{2} \frac{d(V_G - v_L)}{dt} \tag{8.47}$$

and

$$\frac{v_S}{R_S} + C_S \frac{dv_S}{dt} = i_d + \frac{C_G}{2} \frac{d(V_G - v_S)}{dt} \tag{8.48}$$

where v_L and v_S are the error voltages at the data-holding node and the signal-source node, respectively. Gate voltage is assumed to decrease linearly with time from the turn-on value V_H:

$$V_G = V_H - \alpha t \tag{8.49}$$

where α is the falling rate. When the transistor is biased in the strong inversion region,

$$i_d = \beta(V_{HT} - \alpha \cdot t)(v_L - v_S) \tag{8.50}$$

where

$$\beta = \mu C_{OX} \frac{W}{L} \tag{8.51}$$

and

$$V_{HT} = V_H - V_S - V_{thn} \tag{8.52}$$

Here, V_{thn} is the transistor effective threshold voltage, including the body effect. For small-geometry transistors, narrow- and short-channel effects should be considered in determining the V_{thn} value. Under the condition $|dV_G/dt| \gg |dV_L/dt|$ and $|dV_S/dt|$, and Equations 8.47 and 8.48 can be simplified to

$$C_L \frac{dv_L}{dt} = -\beta(V_{HT} - \alpha t)(v_L - v_S) - \frac{C_G}{2}\alpha \tag{8.53}$$

and

$$\frac{v_S}{R_S} + C_S \frac{dv_S}{dt} = \beta(V_{HT} - \alpha t)(v_L - v_S) + \frac{C_G}{2}\alpha \tag{8.54}$$

No closed-form solution to this set of equations can be found. Numerical integration can be employed to find final results. Analytical solutions to special cases are given next.

Figure 8.45a is the circuit schematic diagram of the case, with only a voltage sourced at the signal-source node. Because $C_S \gg C_L$, v_S can be approximated as zero and the governing equation reduces to

$$C_L \frac{dv_L}{dt} = -\beta(V_{HT} - \alpha t)v_L - \frac{C_G}{2}\alpha \tag{8.55}$$

When the gate voltage reaches the threshold condition, the error voltage at the data-holding node is

$$v_L = -\sqrt{\frac{\pi\alpha C_L}{2\beta}}\left(\frac{C_G}{2C_L}\right)\text{erf}\left(\sqrt{\frac{\beta}{2\alpha C_L}}V_{HT}\right) \tag{8.56}$$

Notice that the value of the error function erf(·) can be found from mathematical tables.

Another special case is when the source capacitance is negligibly small, as is shown in Figure 8.45b. The governing equations reduce to

$$C_L \frac{dv_L}{dt} = -\beta(V_{HT} - \alpha t)(v_L - v_S) - \frac{C_G}{2}\alpha \tag{8.57}$$

and

$$\frac{v_S}{R_S} = \beta(V_{HT} - \alpha t)(v_L - v_S) + \frac{C_G}{2}\alpha \tag{8.58}$$

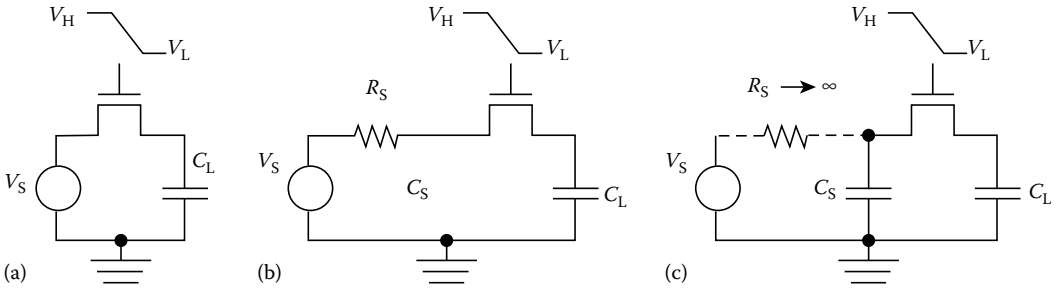


FIGURE 8.45 Special cases of switch charge injection: (a) no source resistance and capacitance; (b) no source capacitance; and (c) infinitely large source resistance.

When the gate voltage reaches the threshold condition, the error voltage at the data-holding node is

$$v_L = -\frac{\alpha C_G}{2C_L}\exp\left(-\frac{V_{HT}}{\alpha C_L R_S}\right)\cdot\int_0^{V_{HT}/\alpha}[\beta R_S(V_{HT}-\alpha\theta)+1]^{1/C_L\beta R_S^2\alpha}$$
$$\cdot\exp\left(\frac{\theta}{C_L R_S}\right)\left(2-\frac{1}{1+\beta R_S(V_{HT}-\alpha\theta)}\right)d\theta \tag{8.59}$$

If a time constant $R_S C_S$ is much larger than the switch turn-off time, then the channel charge will be shared between C_S and C_L, as shown in Figure 8.45c. For the case of a symmetrical transistor and $C_S = C_L$, half of the channel charge will be deposited to each capacitor. Otherwise the following equations can be used to find the results:

$$C_L\frac{dv_L}{dt} = -\beta(V_{HT}-\alpha t)(v_L - v_S) - \frac{C_G}{2}\alpha \tag{8.60}$$

and

$$C_S\frac{dv_S}{dt} = \beta(V_{HT}-\alpha t)(v_L - v_S) + \frac{C_G}{2}\alpha \tag{8.61}$$

We can multiply Equation 8.61 by the ratio C_L/C_S and then subtract the result from Equation 8.60 to obtain

$$C_L\frac{d(v_L - v_S)}{dt} = -\beta(V_{HT}-\alpha t)\left(1+\frac{C_L}{C_S}\right)(v_L - v_S) - \frac{\alpha C_G}{2}\left(1-\frac{C_L}{C_S}\right) \tag{8.62}$$

When the gate voltage reaches the threshold condition, the amount of voltage difference between the data-holding node and the signal-source node becomes

$$v_L - v_S = -\sqrt{\frac{\pi\alpha C_L}{2\beta(1+C_L/C_S)}}\left(\frac{C_G(1-C_L/C_S)}{2C_L}\right)$$
$$\cdot\,\mathrm{erf}\left(\sqrt{\frac{\beta(1+C_L/C_S)}{2\alpha C_L}}V_{HT}\right) \tag{8.63}$$

References

1. N. Weste and K. Eshraghian, *Principles of CMOS VLSI Design*, 2nd ed. Reading, MA: Addison-Wesley, 1993.
2. J. P. Uyemura, *Fundamentals of MOS Digital Integrated Circuits*, Reading, MA: Addison-Wesley, 1988.
3. D. Radhakrishnan, S. R. Whitaker, and G. K. Maki, Formal design procedures for pass transistor switching circuits, *IEEE J. Solid State Circuits*, 20(2), 531–536, April 1985.
4. Y. Suzuki, K. Odagawa, and T. Abe, Clocked CMOS calculator circuitry, *IEEE J. Solid State Circuits*, 8(6), 734–739, Dec. 1973.
5. K. Yano, T. Yamanaka, T. Nishida, M. Saito, K. Shimohigashi, and A. Shimizu, A 3.8-ns CMOS 16×16-b multiplier using complementary pass-transistor logic, *IEEE J. Solid State Circuits*, 25(2), 388–395, April 1990.

6. Y. Oowaki et al., A 7.4ns CMOS 16 × 16 multiplier, in *IEEE: International Solid-State Circuits Conference Digest of Technical Papers*, San Francisco, California, 1987.
7. A. P. Chandrakasan, S. Sheng, and R.W. Brodersen, Low-power CMOS digital design, *IEEE J. Solid-State Circuits*, 27(4), 473–484, April 1992.
8. K. Shimohigashi and K. Seki, Low-voltage ULSI design, *IEEE J. Solid-State Circuits*, 28(4), 408–413, April 1993.
9. P. G. Gray and R. G. Meyer, MOS operational amplifier design—A tutorial overview, *IEEE J. Solid-State Circuits*, 17(6), 969–982, Dec. 1982.
10. R. Gregorian and G. C. Temes, *Analog MOS Integrated Circuits for Signal Processing*, New York: John Wiley & Sons, 1986.
11. D. M. Pietruszynski, J. M. Steininger, and E. J. Swanson, A 50-Mbit/s CMOS monolithic optical receiver, *IEEE J. Solid-State Circuits*, 23(6), 1426–1433, Dec. 1988.
12. G. Williams, U.S. Patent, 4,574,249, Mar. 4, 1986.
13. P. K. Simpson, Foundations of neural networks, in *Artificial Neural Networks: Paradigms, Applications, and Hardware Implementations*, E. Sánchez-Sinencia and C. Lau, Eds. New York: IEEE Press, 1992, pp. 3–24.
14. S. M. Gowda, B. J. Sheu, J. Choi, C.-G. Hwang, and J. S. Cable, Design and characterization of analog VLSI neural network modules, *IEEE J. Solid-State Circuits*, 28(3), 301–313, Mar. 1993.
15. M. Ismail, S. V. Smith, and R. G. Beale, A new MOSFET-C universal filter structure for VLSI, *IEEE J. Solid-State Circuits*, 23(2), 183–194, Feb. 1988.
16. R. E. Geiger, P. E. Allen, and N. R. Strader, *VLSI Design Techniques for Analog and Digital Circuits*, New York: McGraw-Hill, 1990.
17. M. Banu and Y. Tsividis, Fully integrated active RC filters in MOS technology, *IEEE J. Solid-State Circuits*, 18(6), 644–651, Dec. 1983.

9

Digital Systems

9.1 Programmable Logic Devices **9**-1
PLD Device Technologies • PLD Notation • Programmable Logic Array • Programmable Read Only Memory • Programmable Array Logic • Classification of Combinational Logic PLD Devices • Designing with Combinational Logic PAL Devices • Designing with Sequential PAL Devices • Designing with PALs Having Programmable Macrocell Outputs • FPGA Technologies • FPGA Architectures • Design Process • VHDL Synthesis Style for FPGAs • Synthesis of State Machines
References **9**-32
9.2 Clocking Schemes **9**-32
Introduction • Clocking Principles • Clock Distribution Schemes • Future Directions
References **9**-70
9.3 MOS Storage Circuits **9**-72
Dynamic Charge Storage • Shift Register • Dynamic CMOS Logic
References **9**-82
9.4 Microprocessor-Based Design **9**-82
Introduction • Features of a Microprocessor-Based System • Memory • Microprocessor Architecture • Design with a General Purpose Microprocessor • Interfacing • Design with a Microcontroller • Design Guidelines
References **9**-111
9.5 Systolic Arrays **9**-111
Concurrency, Parallelism, Pipelining, and Systolic Array • Digital Filters • Systolic Word and Bit-Level Designs • Recursive LSs Estimation • Kalman Filtering • Eigenvalue and SVDs
References **9**-142

Festus Gail Gray
Virginia Polytechnic Institute and State University

Wayne D. Grover
University of Alberta

Josephine C. Chang
University of Southern California

Bing J. Sheu
Taiwan Semiconductor Manufacturing Co.

Roland Priemer
University of Illinois at Chicago

Kung Yao
University of Southern California, Los Angeles

Flavio Lorenzelli
University of Milan, Crema

9.1 Programmable Logic Devices

Festus Gail Gray

Traditional programmable logic devices (PLDs) and field programmable gate arrays (FPGAs) allow circuit designers to implement logic circuits with fewer chips relative to standard gate level designs based on primitive gates and flip-flops. As a result, layout and unit production costs are generally reduced. In this chapter, we use the term ''programmable device'' to refer to the class of moderately complex single-chip devices, in which the user in the field can program the function of the device. We include

TABLE 9.1 Complexity Ladder of Devices

Device	Complexity of a Single Chip	Range of Realizable Functions	Initial Design and Cost of Design Change	Unit Product Cost (High Volume)	Time to Market
SSI discrete gate chip	Lowest	All functions	High	High	High
MSI chip		Narrow	High	High	High
PLD[a]		Moderate	Low	Moderate	Low
LSI (PROM)[a]		All combinational functions with n inputs	Very low	Moderate	Low
FPGA		Wide	Low	Moderate	Low
MPGA		Wide	High	Low	High
Standard cell		Wide	High	Low	High
Custom chip	Highest	All functions	Very high	Very low	Very high

[a] Field programmable.

such devices as the programmable logic array (PLA), programmable array logic (PAL), programmable read-only memory (PROM), and the FPGA. Since most commercial vendors provide software design aids for mapping designs to their specific chips, initial design costs and time to market are low. Another important advantage of programmable device designs is ''flexibility.'' Design changes do not require physical changes to the printed circuit board as long as the revised functions still fit onto the same programmable chip. The low cost of design revisions makes programmable chips very attractive for prototype design and low volume production. Designers often move up the design ladder once proven designs move into high volume production.

Table 9.1 shows the position of PLDs, PROMs, and FPGAs on the complexity ladder of device types. In the ''range of realizable functions'' column, we compare the range of realizations for various device types. Discrete gates can implement any function if enough gates are available. MSI chips implement very specialized functions such as shift registers, multiplexers (MUXs), decoders, etc. Table 9.1 compares programmable devices (PLDs, FPGAs, and PROMs) relative to the range of functions that can be implemented on a single chip. A PROM chip with n address inputs can implement any combinational function of n variables. A PLD chip with n inputs can implement only a subset of the combinational functions of n variables. Gate arrays can implement a wide range of both combinational and sequential functions. The programmable devices are characterized by low time to market, low design cost, low cost of modifications, and moderate production costs. Nonfield programmable devices such as mask programmable gate arrays (MPGAs), standard cell devices, and full custom devices are characterized by high initial design costs and longer time to market, but have lower volume production costs. Custom chips are preferred for large volume production because of the very low unit production costs. However, initial design costs and the cost of design changes are very high for custom chip design. Also, the design of custom chips requires highly trained personnel and a large investment in equipment. The low design cost, low cost of design changes, and low time to market make PLDs and FPGAs good choices for lower volume production and for prototype development.

The primary difference between PLDs and FPGAs arise because of a difference in the ratio of the number of combinational logic (CL) gates to the number of flip-flops. PLD devices are better for applications that require complex CL functions that drive a relatively small number of flip-flops, such as finite-state machine (FSM) controllers or purely CL functions. FPGAs are better for devices that require arithmetic operations (adders, multipliers, and arithmetic logic units [ALUs]), or that require a large number of registers and less complex CL functions, such as digital filters.

9.1.1 PLD Device Technologies

Companies produce PLD devices in different technologies to meet varying design and market demands. There are two categories of technologies. ''Process technology'' refers to the underlying semiconductor

structure, which affects device speed, power consumption, device density, and cost. ''Programming technology'' refers to the physics of chip programming and affects ease of programming and the ability to reprogram or to reconfigure chips.

9.1.1.1 Process Technologies

The dominant technologies in PLD devices are bipolar and CMOS. Bipolar devices are faster and less expensive to manufacture, but consume more power than CMOS devices. The higher power requirements of bipolar devices limit the gate density. Typical CMOS devices, therefore, achieve much higher gate densities than bipolar devices. The power consumption of CMOS devices depends on the application because a CMOS device only consumes power when it is switching states. The amount of power consumed increases with the speed of switching. Therefore, the total amount of power consumed depends on the frequency and speed of state changes in the device. Some devices have programmable power standby activation that puts the device in a lower power consumption configuration if no input signal changes for a predefined amount of time. The device then responds to the next input change much slower than normal but switches back to the faster speed configuration and maintains the faster speed as long as input changes continue to occur frequently. When programmed in the ''standby power mode,'' power consumption is reduced on the average at the expense of response time of the device. When programmed to operate in the ''turbo'' mode, the device stays in the faster configuration at all times. The result is higher power consumption, but faster response time. The mode of operation is selected by the user to match the requirements of an application.

To take advantage of the higher densities of CMOS devices and still be compatible with bipolar devices, many CMOS PLAs have special driver circuits at the input and output pins to allow pin compatibility with popular bipolar devices such as the commonly used TTL devices.

ECL is a very high-speed technology used in some PLD devices. Although ECL has the highest speed of the popular technologies, the power consumption is very high which severely limits the gate density.

Security is another issue that is related to process technology. Many PLDs have a programmable option that prevents reading the program. Since the software provided by most manufacturers allows the user to read the program in the chip in order to verify correct programming, it is extremely easy to copy designs. To prevent illegal copying of patented designs, one simply blows the ''security fuse,'' which permanently prevents anyone from reading the program by normal means. However, the program in most bipolar circuits can easily be read by removing the case and examining the programmed fuses under a microscope. CMOS PLDs are much more secure because it is virtually impossible to determine the program by examining the circuit.

9.1.1.2 Programming Technologies

The programming technologies used in PLDs are virtually the same as the programming technologies available for read-only memories (ROMs). Programming technologies are divided into two broad categories: mask programmable devices and field programmable devices.

In mask programmable technologies, identical base chips are produced in mass. The final metallization step is simply omitted. A mask programmable PLD chip is programmed by performing a final metal deposition step that selects the programming options. Clearly, this step must be performed at the manufacturer's plant. The user makes entries on order forms that specify how the chip is to be programmed and sends it to the manufacturer. The manufacturer must then prepare one or more production masks prior to making the chip. Mask programmable devices incur a high setup cost to make the first device, but unit production costs are typically less than half of that for field programmable technologies. The usual practice is to use field programmable devices for prototype work and implement only proven designs in mask programmable technologies when a large production volume is required. Many PLD devices are available in both mask programmable and field programmable versions, which make the conversion easy and reliable.

The user can program field programmable technologies directly. Specialized equipment is needed. Modern programming devices can actually program both ROM and PLD devices. The programmer is typically controlled by a small computer (PC) and uses files prepared in standard format (JEDEC) by software provided by the manufacturer or written by software vendors. Such software can include elegant features such as a programming language (ABEL, VHDL, Verilog, etc.), truth table input, equation input, or state machine input. Selection of a chip vendor should include careful evaluation of the support software for programming the chip.

Field programmable PLD technologies can be classified into three broad categories: fusible link PLDs, ultraviolet erasable PLDs (EPLDs), and electrically erasable PLDs (EEPLDs). Field programmable ROMS come in analogous forms: fusible link ROMs (PROMs), ultraviolet erasable ROMS (EPROMs), and electrically erasable ROMS (EEPROMs).

Fusible link PLDs typically utilize bipolar process technology. The programmer blows selected fuses in the device. Because higher than normal voltages and currents are required to blow the fuses, programming fusible link PLDs can be quite stressful for the device. Overheating is a common problem. However, this technology is quite well developed and the design of programming devices is sufficiently mature so that reliable results can be expected as long as directions are carefully followed. Fusible link technologies provide the convenience of on-site programming, which reduces the time required to develop designs and the time required to make design changes. The trade-off involves at least a twofold increase in per unit cost and a significant reduction in device density relative to mask programmable devices because the fuses take up considerable chip space. A fusible link PLD can be programmed only once because the blown fuses cannot be restored.

''Ultraviolet EPLDs'' have a window on the top of the chip. Programming the chip involves storing charges at internal points in the circuit that control switch settings. Shining ultraviolet light through the window on the chip can dissipate the charges. Therefore, EPLDs provide the convenience of reprogramming as a design evolves. On the downside, EPLDs cost at least three times as much per chip as mask programmable PLDs and operate at much slower speeds. Since EPLDs typically utilize CMOS technology, they are slower than fusible link PLDs, but require less power. Therefore, EPLDs are often used in development work with the final design being implemented in either fusible link technology (for faster speed) or mask programmable technology (for faster speed and lower density). In spite of the fact that EPLDs cost more than fusible link PLDs, the reprogramming feature eventually results in a lower cost for development than using fusible link PLDs. This technology requires an additional piece of hardware to erase the chips.

EEPLDs provide the convenience of reprogramming without the need to erase the previous program because the chip is programmed by setting the states of flip-flops inside the device. It is, therefore, not necessary to purchase an erasing device. The reprogramming also requires less time to accomplish. Of course, EEPLD chips cost more and have a lower gate density than EPLD chips.

9.1.2 PLD Notation

PLDs typically have many logic gates with a large number of inputs. Also, there are often many gates that have the same set of inputs. For example, the PAL22V10 has 132 AND gates, each with the same 44 gate inputs. Obviously, the diagram for such a complex circuit using standard AND gate symbols would be extremely complex and difficult to read.

Figure 9.1 is the conventional diagram for an eight-input AND gate. Clearly, a similar diagram for a 44-input AND gate would be very cumbersome. Figure 9.2 is the same eight-input AND gate in PLD notation. The eight parallel

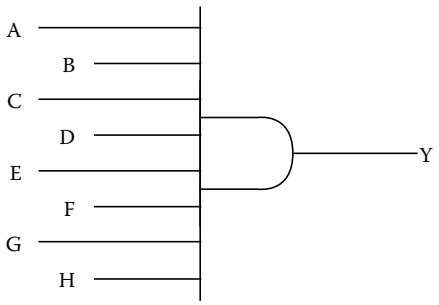


FIGURE 9.1 Conventional diagram for an eight-input AND gate.

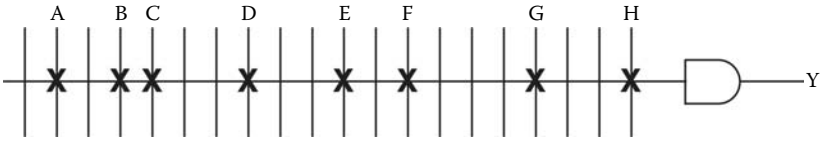


FIGURE 9.2 PLD notation for an eight-input AND gate.

wires that actually occur as inputs to the AND gate are represented by a single horizontal line in PLD notation. The actual inputs to the AND gate are drawn perpendicular to the single line. There are usually more signal lines than just the eight needed for this gate. An *X* is placed at the intersection of the single line with each of the perpendicular lines that provide actual inputs to the AND gate. Keep in mind that the single horizontal line actually represents eight parallel wires that are not physically connected to each other.

By comparing the internal structures of PLAs, PALs, and PROMs, we will describe the capabilities and limitations of each type of PLD. Since the PAL is currently the most popular PLD device, we will describe design methodology for both combinational and sequential PAL devices. By emphasizing the difference between designing with PALs and designing with standard logic gates, we provide practical insights about PLD design.

9.1.3 Programmable Logic Array

Figure 9.3 shows that the basic PLA consists of a programmable AND array followed by a programmable OR array. Vertical lines in the AND array represent the input variables (A, B, C, D). Since each input drives many AND gates, an internal buffer provides high current drive signals in both true and complemented format for each input variable. Initially, there is a connection from each input variable and its complement to each AND gate. In this example circuit, each AND gate initially has eight inputs ($A, \overline{A}, B, \overline{B}, C, \overline{C}, D, \overline{D}$). Each AND gate input line contains a fuse or electronic switch. We program the chip by blowing the fuses in lines that we do not need, or by programming the electronic switches. After

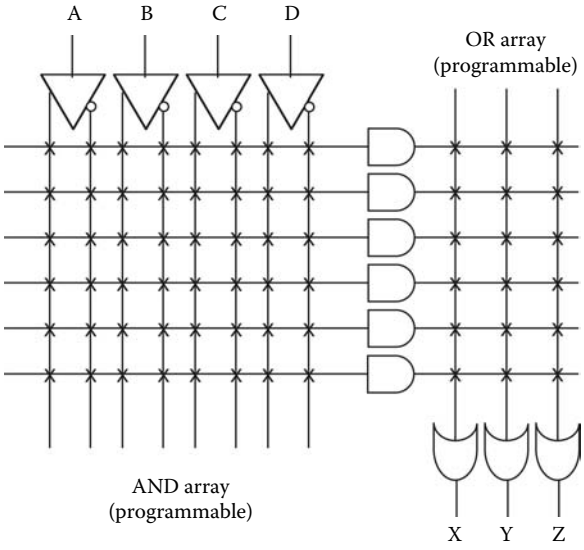
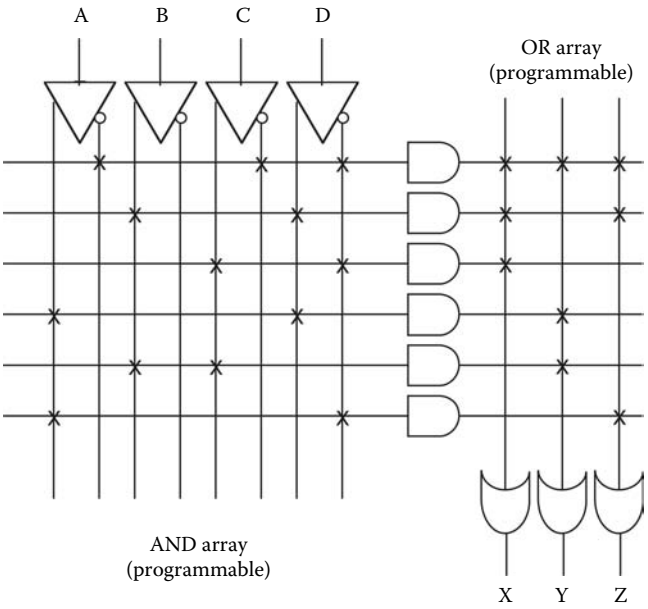


FIGURE 9.3 Basic architecture for a PLA.



$X = \overline{A} \cdot \overline{C} \cdot \overline{D} + B \cdot D + C \cdot \overline{D}$
$Y = \overline{A} \cdot \overline{C} \cdot \overline{D} + A \cdot D + B \cdot C$
$Z = \overline{A} \cdot \overline{C} \cdot \overline{D} + B \cdot D + A \cdot \overline{D}$

FIGURE 9.4 An example of a programmed PLA.

programming, we remove the X's from the lines that are disconnected. For example, in the programmed chip of Figure 9.4, the upper AND gate implements product term $(\overline{A} \cdot \overline{C} \cdot \overline{D})$.

In the OR array of Figure 9.3, there is an initial connection from each AND gate output to every input on each OR gate. Again, the single vertical line connected on the input side of each OR gate represents all six wires. Each of the input lines to the OR gates also contains a fuse or programmable switch. Figure 9.4 shows that, after programming, output X connects to product terms $\overline{A} \cdot \overline{C} \cdot \overline{D}$, $B \cdot D$, and $C \cdot \overline{D}$.

The number of product lines on a chip limits the range of functions that fit onto the chip. The PLA chip in Figure 9.3 can implement any three functions (X, Y, Z) of the same four variables (A, B, C, D) as long as the total number of required product terms is less than or equal to six. However, there are 100 different product terms involving four variables. So, practical chips have many more product lines than this contrived example.

In order to fit functions onto the chip, designers must be able to simplify multiple output functions using gate sharing whenever possible. Finding a minimal gate implementation of multiple output functions with gate sharing is a very complex task. The goal is to minimize the total number of gates used. The size of gates does not matter. For example, whether an AND gate has four inputs or two inputs is not important. All that changes are the number of fuses that are blown. This differs dramatically from the minimization goals when discrete gates are used. For discrete gate minimization, a four-input gate costs more than a two-input gate. Therefore, the classical minimization programs need to be modified to reflect the different goals for PLA development.

Three parameters determine the capacity of a PLA chip. Let n be the number of inputs, p the number of product term lines, and m the number of outputs. Then, the PLA chip can implement any m functions of the same n variables that require a total of p or fewer product terms. The device complexity is proportional to $(m + n)p$.

9.1.4 Programmable Read Only Memory

The PROM is the most general of the CL PLD devices described in this section. However, from a structural viewpoint, the PROM is a special case of the PLA in which the AND array is fixed and the OR array is programmable. Figure 9.5 is a conceptual diagram of a PROM. The filled circles in the AND array represent permanent connections. The X's in the OR array indicate that it is programmable. The number of product lines in a PROM is 2^n; whereas the number of product lines in a typical PLA is much smaller. A PROM has a product line for each combination of input variables. Since any logic function of n variables can be expressed in a canonical sum of minterms form in which each product term is a product of exactly n literals, the PROM can implement any function of its n input variables.

To demonstrate the generality of the PROM, Figure 9.6 shows how the PROM of Figure 9.5 must be programmed so as to implement the same set of logic functions that are programmed into the PLA of Figure 9.4. The PROM program follows directly from the truth table for a logic function. The truth table for the logic functions X, Y, and Z appears in Table 9.2. The correspondence between the truth table and the program in the PROM of Figure 9.6 is straightforward. A logic 1 in the truth table corresponds to an X in the figure and a logic 0 in the table corresponds to the absence of an X.

A PROM with n address lines (serving as n input variable lines) and m data lines (serving as m output variable lines) can implement any m functions of the same n variables. Unlike a PLA, a PROM has no restrictions due to a limited number of product lines. The PROM contains an n input, 2^n output decoder that generates 2^n internal address lines that serve as product lines. Since the decoder grows exponentially in size with n, the cost of a PROM also increases rapidly with n. The justification for a PLA is to reduce the cost of the PROM decoder by providing fewer product terms since many practical functions require significantly fewer than 2^n product terms. As a result, some n variable functions will not fit onto a PLA chip with n input variables, whereas any n variable function will fit onto a PROM with n address lines.

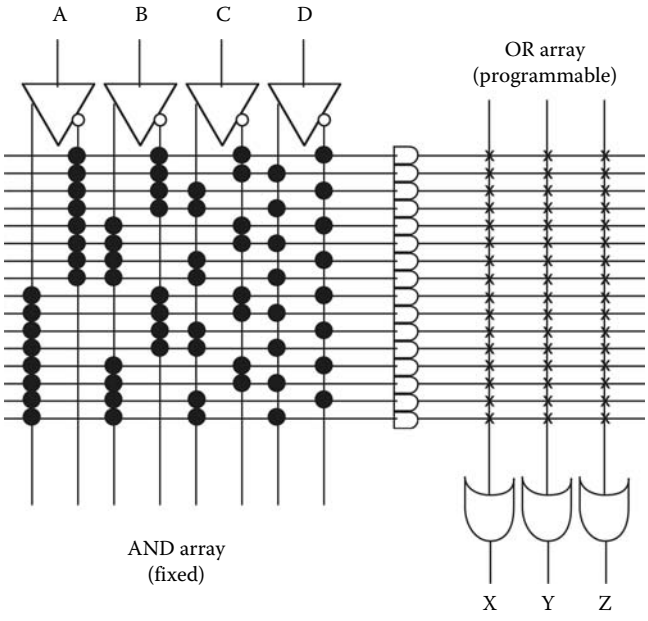


FIGURE 9.5 Conceptual diagram of a PROM.

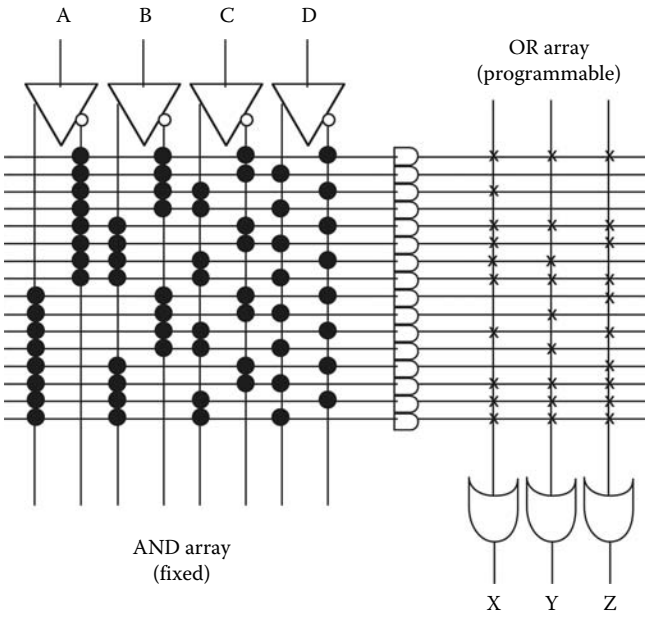


FIGURE 9.6 An example of a programmed PROM.

TABLE 9.2 Truth Table for the Logic Functions Implemented in the PROM

ABCD	XYZ	ABCD	XYZ
0000	111	1000	001
0001	000	1001	010
0010	100	1010	101
0011	000	1011	010
0100	111	1100	001
0101	101	1101	111
0110	110	1110	111
0111	111	1111	111

9.1.5 Programmable Array Logic

PAL is the most popular form of PLD today. Lower price, higher gate densities, and ease of programming all tend to make PAL more popular than PLA. On the negative side, the range of functions that can fit onto a chip with the same number of inputs, outputs, and product lines is less for a PAL than for a PLA.

Figure 9.7 is the basic architecture of a PAL. The PAL architecture is a special case of the PLA architecture in which the OR array is fixed. The filled circles in the OR array indicate permanent connections. Only the AND array is programmable. Compare this PAL architecture with the PLA architecture in Figure 9.3. Since the OR array is not programmable, it is immediately evident that fewer functions will fit onto the PAL. In the PLA, the product terms can be divided among the three

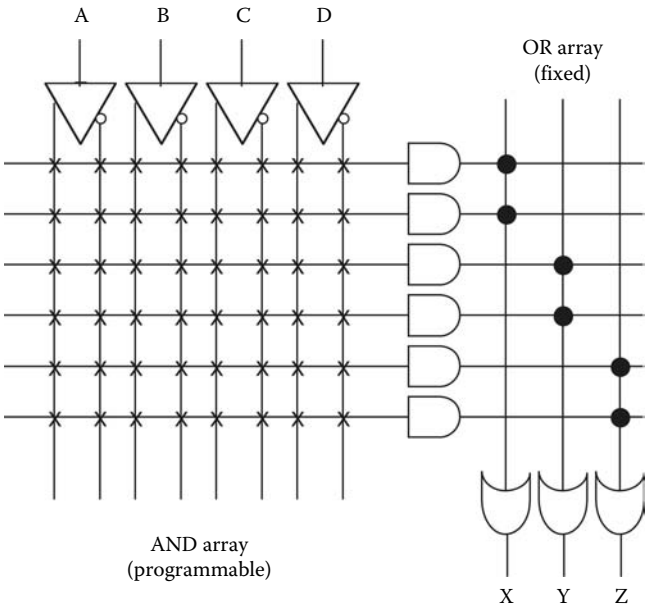


FIGURE 9.7 Basic architecture of a PAL.

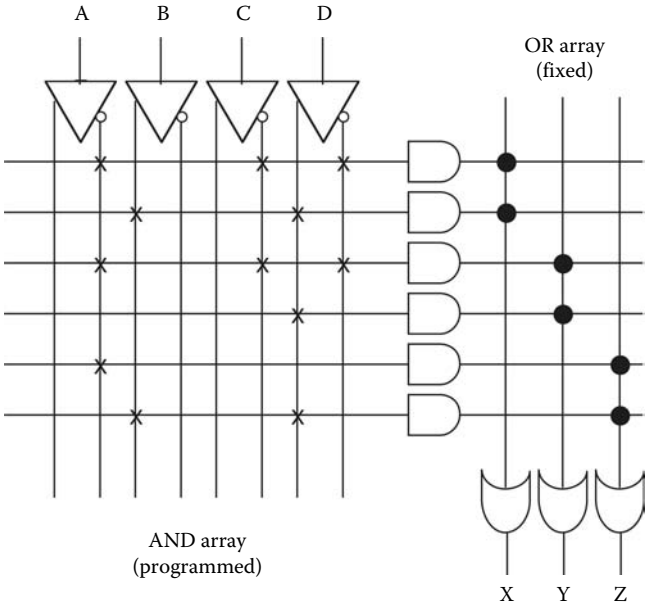
outputs in any way desired and product terms that are used in more than one output can share the same product line. In the PAL, each output is limited to a fixed number of product terms. In Figure 9.7, all outputs are limited to two product terms. In addition, if two output functions both require the same product term in a PAL, two different product lines must be used.

Consider the three functions implemented on the PLA in Figure 9.4. Since the three functions require a total of nine product terms, they will not fit onto the PAL of Figure 9.7. However, any function that would fit onto this PAL would obviously fit onto the PLA since the OR array in the PLA can be programmed to be identical to the OR array of the PAL. Figure 9.8 is an example of three functions that fit onto this PAL. Note that we must use two different product lines to provide the same product term ($\overline{A} \cdot \overline{C} \cdot \overline{D}$) to outputs X and Y.

In order to describe the range of applications for a PAL, we must know the number of inputs, n, the number of outputs, m, and the number of product lines that are permanently connected to each output OR gate. The PAL of Figure 9.7 has four inputs, $n = 4$, three outputs, $m = 3$, and has two product lines connected to each OR gate. This PAL is described as a 2-2-2 PAL with four-input variables. Many PALs have the same number of product terms permanently connected to each output. In this case, the three parameters n, m, and p completely describe the size of the PAL. For PALs, the parameter p usually represents the number of product terms per output instead of the total number of product terms, as was the case for PLAs.

The minimization algorithm for multiple output PALs is significantly less complex than the minimization algorithm for a PLA because gate sharing is eliminated as a possibility by the fact that the OR array is not programmable. This means that each output function can be minimized independently. Minimizing a single output function is much less complex than minimizing a set of output functions where gate sharing must be considered.

Overall, the higher density, less complex minimization algorithms, and lower cost of PALs tend to offset the additional functional capabilities of PLAs.



$X = \overline{A} \cdot \overline{C} \cdot \overline{D} + B \cdot D$

$Y = \overline{A} \cdot \overline{C} \cdot \overline{D} + D$

$Z = \overline{A} + B \cdot D$

FIGURE 9.8 An example of a programmed PAL.

9.1.6 Classification of Combinational Logic PLD Devices

The programmability of the AND and OR array provide a convenient means to classify CL PLD types. The classification of Table 9.3 illustrates comparative features of CL PLD devices. Even though PLAs have the most general structure (i.e., both the AND and OR arrays are programmable), the number of functions that fit onto the chips is limited by the number of product terms per output. ROMs and PROMs have fixed AND arrays, but all possible product terms are provided. Therefore, PROMs and ROMs are the most general devices from a functional viewpoint. Applications are only limited by the size and cost of available devices. PALs are the most restricted devices from both the structural and functional viewpoints. Nevertheless, the lower cost relative to PROMS and PLAs, higher gate densities relative to PLAs, and wider variety of available chip types have contributed to a rapid rise in popularity of PAL devices. For this reason, we will concentrate on PAL devices in the rest of this section.

TABLE 9.3 Classification of Combinational PLD Devices

AND Array	OR Array	Device	Typical Number of Product Terms per Output Gate
Fixed	Mask programmable	ROM	2^n
Fixed	Field programmable	PROM, EPROM, EEPROM	2^n
Field programmable	Fixed	PAL	16
Field programmable	Field programmable	PLA	50–150

9.1.7 Designing with Combinational Logic PAL Devices

Determining if a function will fit onto a PAL device is a complex procedure. To demonstrate the difficulty, we will examine the types of functions that can fit onto a PAL16L8 chip. Figure 9.9 shows that the PAL16L8 chip has eight-output pins and 16 pairs of vertical input lines to the AND array. The "L" in the chip name indicates that the eight outputs are all active low CL outputs. The most important information not provided by the device name is the number of product terms per output, which is seven for this device. An additional product term provides an output enable for each output pin, so there are eight product lines per output pin (a total of 64 product lines). There are 10 primary input pins (pin numbers 1–9 and 11). In terms of our definitions, it would seem that $n = 10$, $m = 8$, and $p = 7$. As we will demonstrate, this simplistic analysis significantly understates the capacity of this chip.

A simplistic analysis would say that the PAL16L8 chip could implement any eight functions of the 10 input variables as long as each function requires no more than seven product terms. As far as it goes, this statement is correct. However, it significantly understates the capacity of the chip because it does not take into account the fact that six of the output pins are internally connected as inputs to the AND array (pins 13–18). This is the source of the additional six inputs to the AND array. These internal feedback connections significantly expand the capacity of the chip.

Consider the following logic function.

$$X = A\overline{B}C + B\overline{C}D + \overline{A}E + D\overline{E}F + \overline{A}C + \overline{D} + \overline{F}GH + F\overline{G}I + BE\overline{H} + \overline{C}H + \overline{I}J + \overline{B}E\overline{J} + \overline{D}H$$

It appears that this logic function will not fit onto the PAL16L8 chip because it requires 13 product terms and each output has only seven product lines. However, if not all chip outputs are needed for the application, we can use one of the feedback lines to fit this function onto the chip.

We first partition the function X as follows:

$$X = A\overline{B}C + B\overline{C}D + \overline{A}E + D\overline{E}F + \overline{A}C + \overline{D} + Y$$
$$Y = \overline{F}GH + F\overline{G}I + BE\overline{H} + \overline{C}H + \overline{I}J + \overline{B}E\overline{J} + \overline{D}H$$

Since Y has only seven product terms, we will map function Y to the macrocell connected to pin 18 and use the feedback line from pin 18 to connect Y to a pair of vertical lines in the AND array. Function Y is now available to all other cells as an input variable. Since function X also has seven product terms, we will map X to the macrocell connected to pin 17. One of the product terms in X is the single variable Y. Figure 9.10 shows the final implementation. To obtain the needed product terms for output X, we used two macrocells in the array. As a result, pin 18 is no longer available as an output.

Two practical matters need to be considered. First, some signals must now pass through the AND array twice as they proceed from an input to an output due to the feedback path. Therefore, the delay of the chip is now twice what it was when the feedback path was not utilized. Second, the buffer inverts the outputs; therefore, we actually obtain X on pin 17. If X is specified to be active low, then the output on pin 17 is exactly what we need. If the output X is specified to be active high, then an inverter is required. Since PALs are available with both active low and active high outputs, a designer should select an appropriate PAL to eliminate the need for inverters.

Another feature that adds to the flexibility of the chip is that pins 13–18 can be used either as inputs or as outputs. For example, the enable for the output buffer at pin 15 can be permanently disabled. Since pin 15 is connected directly into the AND array, it is no different from any other input, say pin 2. Of course, by permanently disabling the buffer at pin 15, the OR array connected to the buffer is also disconnected. In order to use pin 15 as an input, we must give up the use of the macrocell connected to pin 15. However, dual use pins and feedback lines dramatically extend the range of applications for the chip. For example, suppose we need only one output and select pin 19 for that use. Then, pins 13–18 are available either as inputs or as feedback lines. We can therefore fit a wide range of functions onto the chip with varying numbers of inputs and product terms.

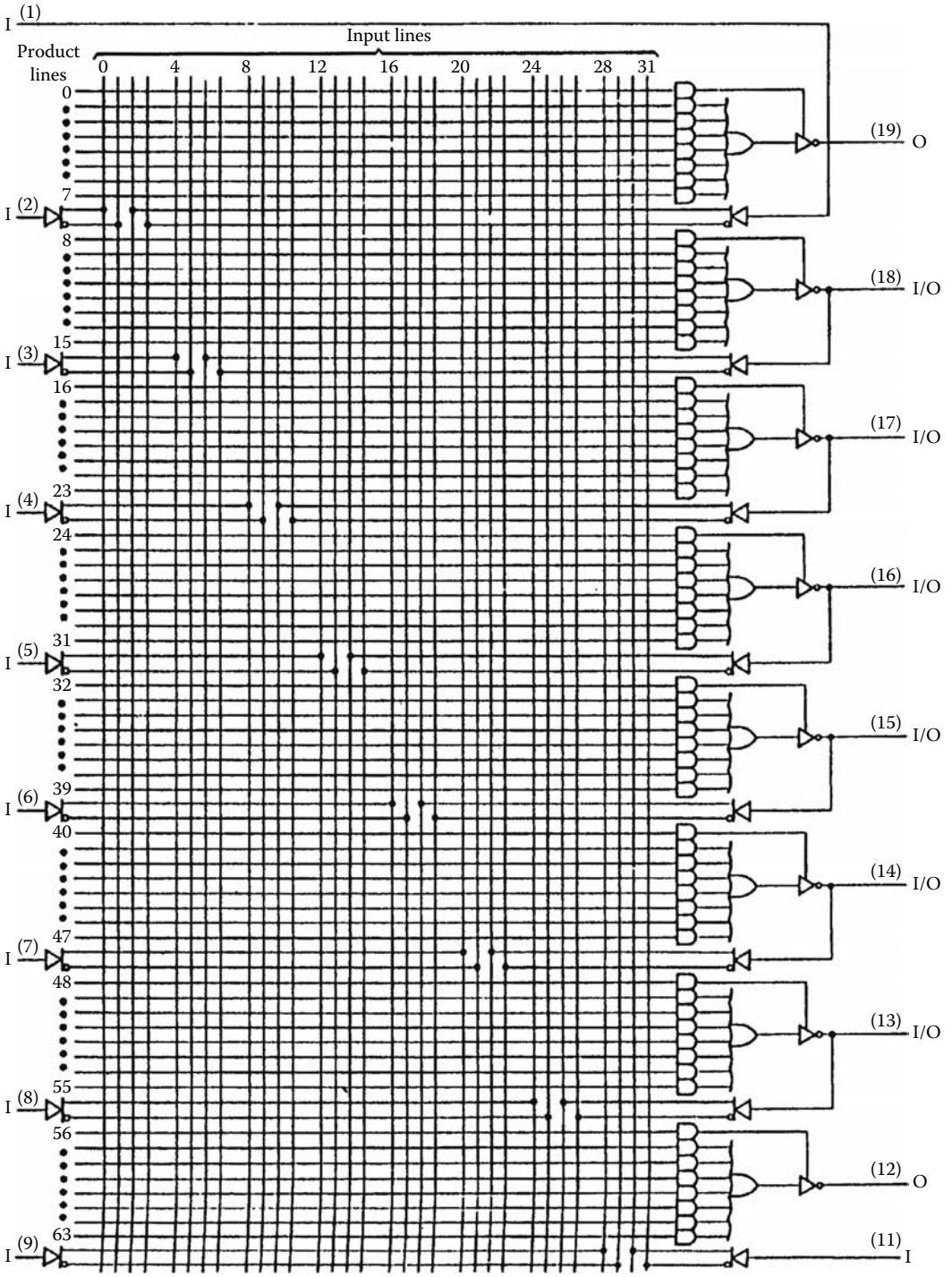


FIGURE 9.9 Logic diagram of PAL 16L8 chip. (Courtesy of Texas Instruments, Dallas.)

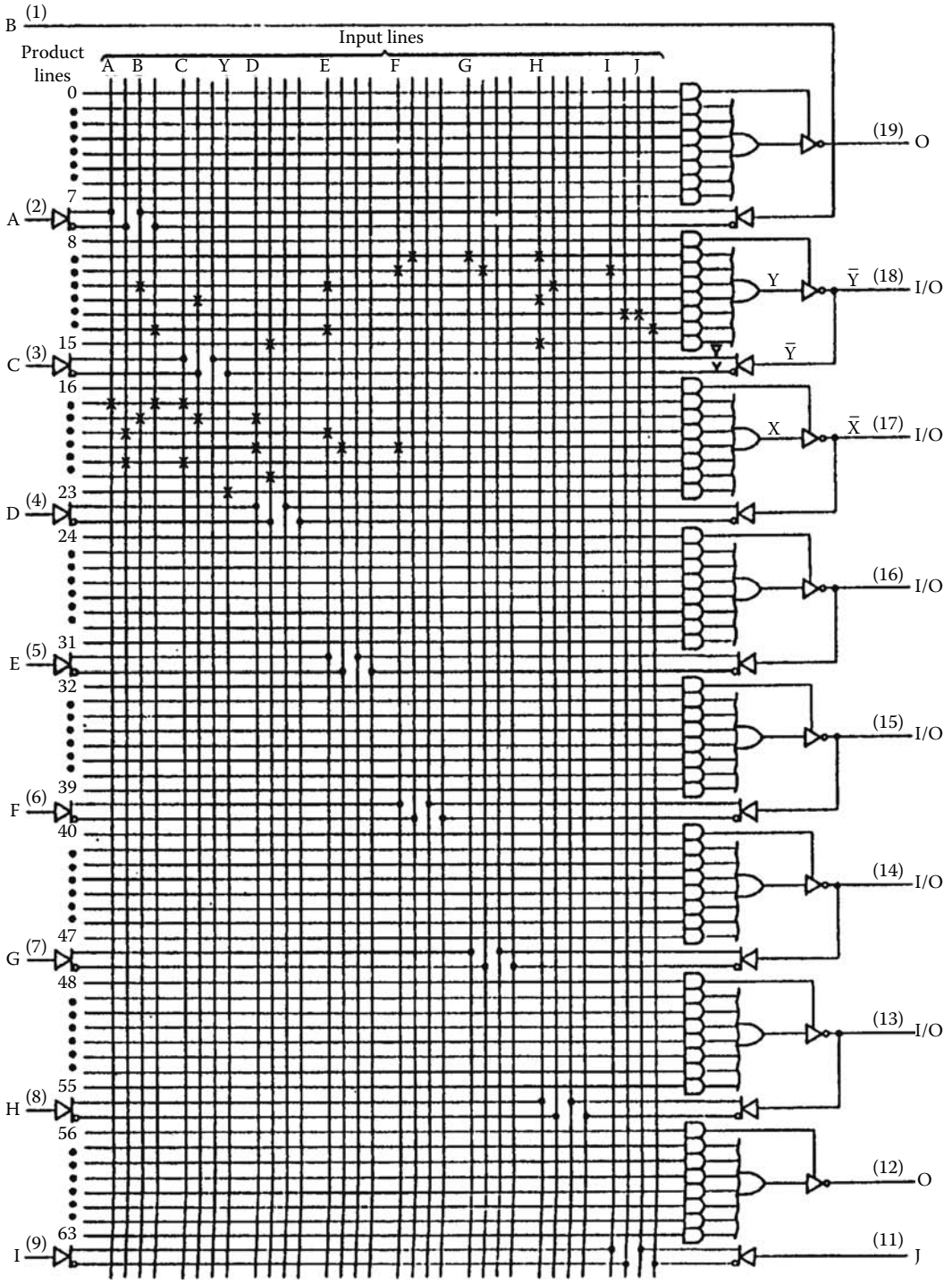


FIGURE 9.10 Implementation of function with 13 product terms on PAL16L8 chip.

Assuming that we need only the one output on pin 19, then we could have up to 16 inputs (pins 19, 11, 13–18). We could, of course, use only the seven product lines in the macrocell connected to pin 19 to implement our function. We, therefore, conclude that the PAL16L8 chip can implement any single function of 16 variables that requires no more than seven product terms.

If we need more product terms but not as many input variables, then we can connect pin 11 to pin 12. This connects the output on pin 12 back into the AND array the same as any other input variable. The output on pin 19 can pick up the seven product terms on pin 12 as an input. This takes up one of the product lines for pin 19, but six product lines remain. Therefore, the single output on pin 19 can now have up to 13 product terms. However, pin 11 is no longer available as an input. Therefore, we must conclude that the PAL16L8 can implement any single function of 15 variables that requires no more than 13 product terms.

If we want to maximize the number of product terms for a single output at pin 19, then we can use pins 13–18 and pin 11 as feedback lines. Each feedback line contributes seven product terms. The AND array for pin 19 can pick up all 49 product terms by using one product line to pick up each feedback variable. All product lines are now busy. The OR gate at pin 19 then sums all 49 product terms. The only pins that are now available for inputs are 1–9. We therefore conclude that the PAL16L8 can implement any single function of nine variables that requires 49 or fewer product terms.

Clearly, there are many combinations of implementations with a variety of values for the number of inputs, the number of outputs, and the number of product terms per output. Table 9.4 shows the full range of possible implementations. For example, from the tables we note that the PAL16L8 can implement any three functions of 10 variables in which the product terms are distributed among the three outputs in any of the following ways: 7–7–37, 7–13–31, 7–19–25, 13–13–25, or 13–19–19. The notation 7–7–37 means that two of the outputs require at most seven product terms and that the third output requires at most 37 product terms. To accomplish these results, five of the output pins must be devoted to feedback lines.

Any implementation that uses a feedback line will have a time delay equal to twice that of a single macrocell. In the delay column of Table 9.4, symbol T_A represents the delay that a signal experiences while passing through the AND–OR array one time. In implementations that do not use feedback lines, signals experience a maximum delay of T_A. For implementations that use one or more feedback lines, the delay is $2T_A$ because some input signals propagate through the AND–OR array twice before reaching an output pin. However, none of the implementations in the table requires more than twice the normal time delay for a single macrocell.

Although the tables cover broad generalizations for classes of functions that will fit onto the PAL16L8 chip, there are certain special types of more complex functions that will fit. For example, suppose that input variables A, B, C, D, E, F, G, H, and I occupy pins 1–9. Further suppose that we implemented functions S, T, V, W, X, Y, and Z using the macrocells connected to pins 12, 13, 14, 15, 16, 17, and 18, respectively. Further suppose that we connect pin 12 to pin 11 so that all of these functions are connected to a pair of vertical input lines in the AND array. Thus, all of these functions are now available to the single output P at pin 19. This approach allows many very complex logic functions to fit onto the chip.

Example 9.1

Let each of the functions S, T, V, W, X, Y, and Z be a sum of products expression involving the nine input variables with at most seven product terms. For example, S might be

$$S = A\bar{B}CD\bar{E}F\bar{G}HI + \bar{A}B\bar{C}DE\bar{F}GH\bar{I} + \bar{B}CD\bar{E}F\bar{G}HI + \bar{F}GH\bar{I} + \bar{A}H\bar{I} + \bar{D}EF\bar{G}HI + ABCDEFGHI$$

Variables T, V, W, X, Y, and Z could be of similar complexity.

TABLE 9.4 Range of Implementations of PAL16L8 Chip

m	*n*	Number of Product Terms per Output	Delay
1	16	7	T_A
1	15	13	$2T_A$
1	14	19	$2T_A$
1	13	25	$2T_A$
1	12	31	$2T_A$
1	11	37	$2T_A$
1	10	43	$2T_A$
1	9	49	$2T_A$
2	16	7–7	T_A
2	15	7–13	$2T_A$
2	14	7–19, 13–13	$2T_A$
2	13	7–25, 13–19	$2T_A$
2	12	7–31, 13–25, 19–19	$2T_A$
2	11	7–37, 13–31, 19–25	$2T_A$
2	10	7–43, 13–37, 19–31, 25–25	$2T_A$
3	15	7–7–7	T_A
3	14	7–7–13	$2T_A$
3	13	7–7–19, 7–13–13	$2T_A$
3	12	7–7–25, 7–13–19, 13–13–13	$2T_A$
3	11	7–7–31, 7–13–25, 7–19–19, 13–13–19	$2T_A$
3	10	7–7–37, 7–13–31, 7–19–25, 13–13–25, 13–19–19	$2T_A$
4	14	7–7–7–7	T_A
4	13	7–7–7–13	$2T_A$
4	12	7–7–7–19, 7–7–13–13	$2T_A$
4	11	7–7–7–25, 7–7–13–19, 7–13–13–13	$2T_A$
4	10	7–7–7–31, 7–7–13–25, 7–7–19–19, 7–13–13–19, 13–13–13–13	$2T_A$
5	13	7–7–7–7–7	T_A
5	12	7–7–7–7–13	$2T_A$
5	11	7–7–7–7–19, 7–7–7–13–13	$2T_A$
6	12	7–7–7–7–7–7	T_A
6	11	7–7–7–7–7–13	$2T_A$
6	10	7–7–7–7–7–19, 7–7–7–7–13–13	$2T_A$
7	11	7–7–7–7–7–7–7	T_A
7	10	7–7–7–7–7–7–13	$2T_A$
8	10	7–7–7–7–7–7–7–7	T_A

Then, output P might be

$$P = A\overline{B}CD\overline{E}FG\overline{H}IS\overline{T}VW\overline{X}Y\overline{Z} + B\overline{C}DEF\overline{G}\overline{S}T\overline{V}WXY\overline{Z} + \cdots$$

where P has at most seven such product terms.

The delay of this implementation is still twice the delay of one basic macrocell.

Example 9.2

This example illustrates embedded factors. Each equation has at most seven product terms involving the listed variables.

$$S = f(A - I) = \overline{A}BC\overline{D}E\overline{F}GH\overline{I} + A\overline{B}\overline{C}DEF\overline{G}HI + \cdots$$

$$T = f(S, A - I) = AB\overline{C}D\overline{E}FG\overline{H}IS + \overline{A}BCD\overline{E}\overline{F}GHI\overline{S} + \cdots$$

$$V = f(S, T, A - I) = \overline{C}DE\overline{F}GH\overline{I}\overline{S}T + BC\overline{D}HIS\overline{T} + \cdots$$

$$W = f(S, T, V, A - I) = \overline{A}BCD\overline{E}F\overline{G}HI\overline{S}T\overline{V} + \overline{D}E\overline{F}HIS\overline{T}V + \cdots$$

$$X = f(S, T, V, W, A - I) = \overline{A}\overline{B}CDE\overline{F}GH\overline{I}ST\overline{V}\overline{W} + \overline{E}T\overline{V}W + \cdots$$

$$Y = f(S, T, V, W, X, A - I) = ABC\overline{D}EFG\overline{H}IS\overline{T}VW\overline{X} + \overline{D}H\overline{I}S\overline{T}V\overline{W}X + \cdots$$

$$Z = f(S, V, T, W, X, Y, A - I) = A\overline{B}CD\overline{E}FGH\overline{I}ST\overline{V}WX\overline{Y} + \overline{F}H\overline{I}WXY + \cdots$$

$$P = f(S, V, T, W, X, Y, Z, A - I) = A\overline{B}CD\overline{E}FG\overline{H}IS\overline{T}VW\overline{X}Y\overline{Z} + B\overline{C}DEF\overline{G}ST\overline{V}WXY\overline{Z} + \cdots$$

The delay of this implementation is eight times the delay of a single macrocell because an input signal change might have to propagate serially through all of the macrocells on its way to the output at pin 19.

These examples demonstrate that very complex functions can fit onto the chip. Determining the optimum way to factor the equations is a very complex issue. Chip manufacturers and third-party vendors provide software packages that aid in the fitting process.

9.1.8 Designing with Sequential PAL Devices

The concept of registered outputs extends the range of PAL devices to include sequential circuits. Figure 9.11 is the logic diagram for the PAL16R4 chip. Again, this chip has 16 pairs of inputs to the AND array. The R4 part of the designation means that the chip has four outputs connected directly to D type flip-flops, i.e., the outputs are registered. Let us add another parameter, k, to designate the number of flip-flops on the chip. An examination of Figure 9.11 indicates that the PAL16R4 also has four combinational outputs with feedback connections to the AND array. These pins are I/O pins because they can also be used as inputs if the OR output to the pins is permanently disabled. All outputs are active low. The chip has eight-input pins. Using our parameter system, the PAL16R4 apparently has $n = 8$, $m = 4$, $k = 4$, $p = 7$ for combinational pin outputs and $p = 8$ for registered pin outputs. However, as for the PAL16L8, these numbers significantly understate the capabilities of this chip.

Since the four registered outputs are also connected back into the AND array, this chip can implement a sequential circuit with the registered outputs serving as state variables. Therefore, this chip can implement any eight-input, four-output, sequential circuit that needs no more than four-state variables (16 states) and no more than seven product terms for each output or eight product terms for each state variable. Separate pins provide an enable for the state variables (pin 11) and a clock for the flip-flops (pin 1). Thus, the state variables are also available at output pins.

By an analysis similar to that used in the previous section, we can utilize the feedback lines to significantly expand the types of circuits that will fit onto the PAL16R4 chip. Table 9.5 shows the range of basic possibilities.

For example, the table indicates that the PAL16R4 chip can implement any single output, eight-input, sequential circuit that requires no more than four-state variables (16 states) and in which the available product terms may be divided among the outputs and state variables in seven different distributions. The notation (7)-(8–8–8–26) means that the single output can have up to seven product terms, that one state variable can have up to 26 product terms, and that the other three state variables can have up to eight product terms each.

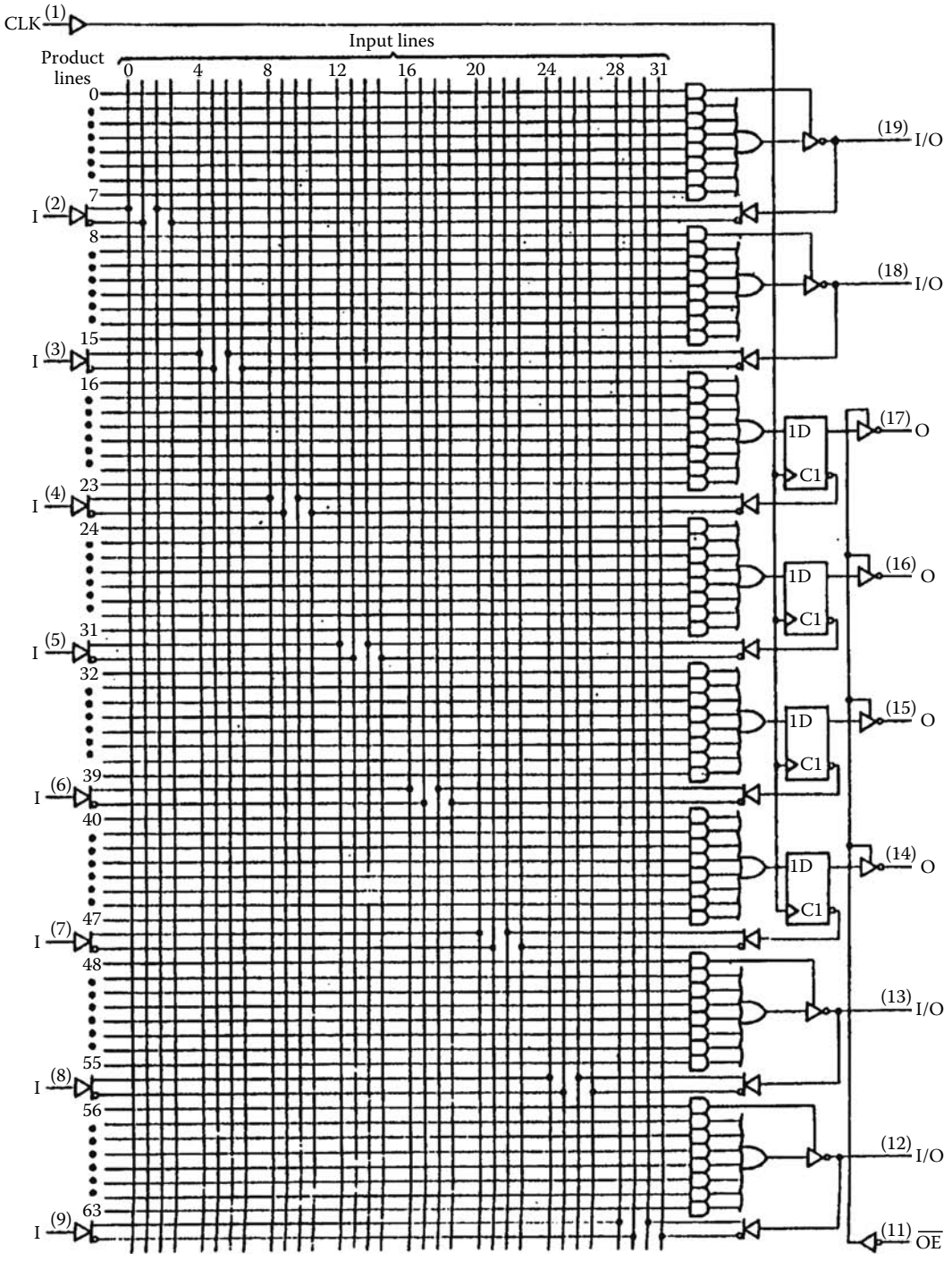


FIGURE 9.11 Logic diagram of PAL16R4 chip.

TABLE 9.5 Range of Basic Implementations of PAL16R4 Chip

M	n	K	Number of Product Terms per (Combinational Output)–(State Variable)	Delay
1	11	4	(7)−(8–8–8–8)	T_A
1	10	4	(7)−(8–8–8–14), (13)−(8–8–8–8)	$2T_A$
1	9	4	(7)−(8–8–8–20), (7)−(8–8–14–14), (13)−(8–8–8–14), (19)−(8–8–8–8)	$2T_A$
1	8	4	(7)−(8–8–8–26), (7)−(8–8–14–20), (7)−(8–14–14–14), (13)−(8–8–8–20), (13)−(8–8–14–14), (19)−(8–8–8–14), (25)−(8–8–8–8)	$2T_A$
2	10	4	(7–7)−(8–8–8–8)	T_A
2	9	4	(7–7)−(8–8–8–14), (7–13)−(8–8–8–8)	$2T_A$
2	8	4	(7–7)−(8–8–8–20), (7–7)−(8–8–14–14), (7–13)−(8–8–8–14), (7–19)−(8–8–8–8), (13–13)−(8–8–8–8)	$2T_A$
3	9	4	(7–7–7)−(8–8–8–8)	T_A
3	8	4	(7–7–7)−(8–8–8–14), (7–7–13)−(8–8–8–8)	$2T_A$
4	8	4	(7–7–7–7)−(8–8–8–8)	T_A

9.1.9 Designing with PALs Having Programmable Macrocell Outputs

The PAL16R4 chip has limited application potential because the outputs from pins 14–17 ''must'' be registered. Most new chips allow a user to decide whether to have registered or combinational outputs at each pin and also allow the user to select either active high or active low outputs.

The PAL22V10 chip (see architecture in Figure 9.12) demonstrates this additional flexibility. Each of 10 macrocells contains a normal PAL AND array and an I/O architecture control block. Each PAL AND array provides a differing number of product terms permanently connected as inputs to an OR gate and an additional product term that enables an output buffer. The number of product terms per output is printed near the OR gate symbol (8, 10, 12, 14, 16, 16, 14, 12, 10, 8). Figure 9.13 shows that this chip is similar in form to the PAL chips described earlier in this chapter. There are 22 vertical pairs of input lines to the AND array. Of these pairs, 11 are connected directly to input pins labeled $I_1 - I_{11}$ (pins 2–11, 13). Ten pairs are feedback lines from the architecture control blocks of the 10 macrocells. Each macrocell is associated with a bidirectional pin (pins 14–23) that can be used either as an input pin, an output pin, or a bidirectional bus pin. If used as a bidirectional bus pin, the designer must control the O/E using a

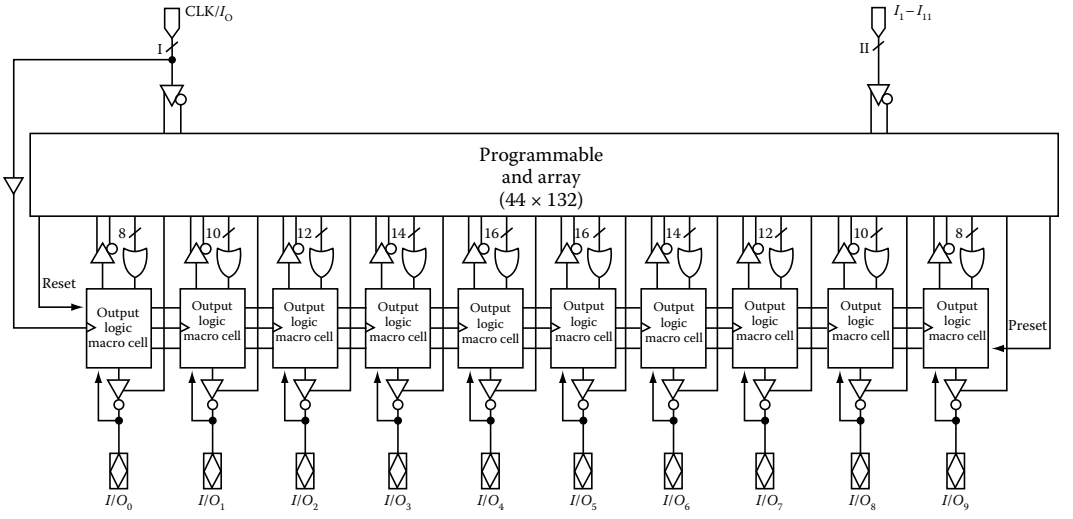


FIGURE 9.12 Architecture of the PAL22V10 chip. (Courtsey of Advanced Micro Devices, Inc., Sunnyvale.)

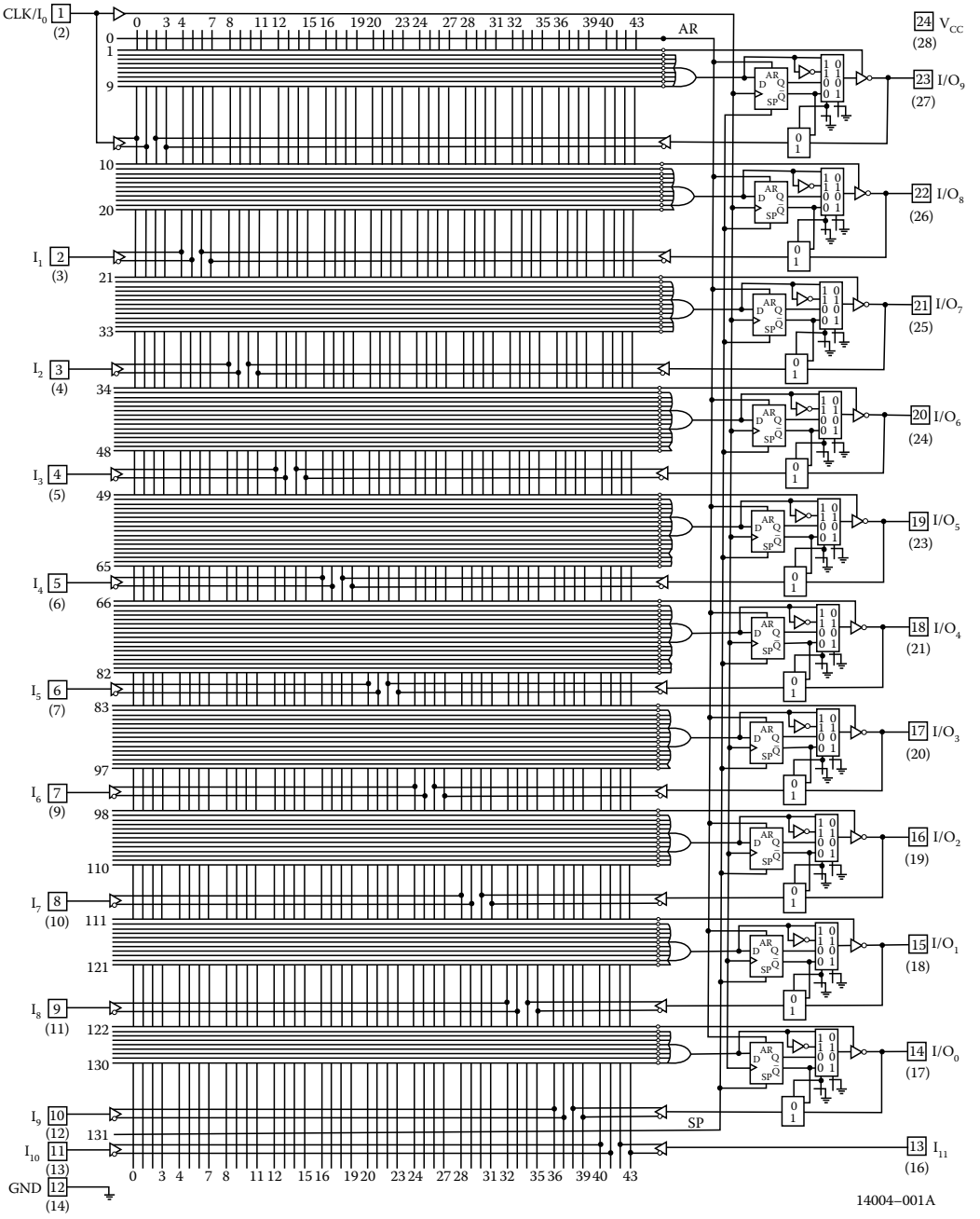


FIGURE 9.13 Complete circuit diagram for the PAL22V10 chip. (Courtesy of Advanced Micro Devices, Inc., Sunnyvale.)

product term from the AND array. The 22nd pair, labeled CLK/I_0, is connected to pin 1. If the chip is being used to implement a purely combinational circuit, pin 1 can be used as an additional input variable. If a sequential circuit is being implemented, pin 1 must be used as the clock signal for the flip-flops.

The architecture control block in each macrocell provides designer control over the signal that is connected to the bidirectional pin and feedback line associated with that macrocell. Figure 9.14 shows

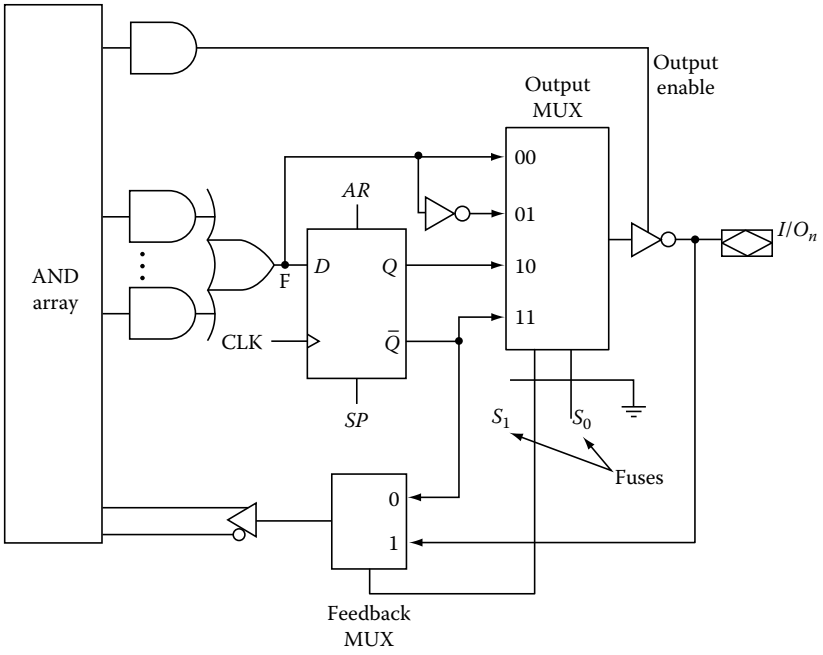


FIGURE 9.14 Macrocell architecture of PAL22V210 chip. (Courtesy of Advanced Micro Devices, Inc., Sunnyvale.)

that the architecture control block contains a D-flip-flop, an inverter, two MUXs with programmable select lines, and an output buffer with an enable. The Output MUX selects either the direct output of the combinational AND array (either active high or active low) or the data value stored in the D-flip-flop (either active high or active low). If the O/E is active, the buffer steers the signal selected by the Output MUX to the pin. An inactive enable causes the buffer to enter the high impedance state, which effectively disconnects the buffer from the pin. The pin can then be an input or can be connected to an external bus.

The feedback signal selected by the Feedback MUX is either the pin signal or the data value stored in the flip-flop (low active). Therefore, the feedback line can be used to expand the number of product terms, to provide a state variable for a sequential circuit, to provide an additional input for the chip, or to implement a bidirectional pin.

Figures 9.13 and 9.14 show that the common clock input to all flip-flops comes directly from pin 1, that a single product (SP) term provides a common synchronous preset for all flip-flops, and that another single product line (AR) provides a common asynchronous reset for all flip-flops. The asynchronous reset occurs when the product line is active independent of clock state. The synchronous preset occurs only on the active edge of the clock when the preset product line is active.

The two programmable MUXs in the architecture control block significantly increase the flexibility of the chip compared to either the PAL16L8 or the PAL16R4. Table 9.6 shows several combinations of switch settings along with typical applications for each setting.

The PAL22V10 is much more versatile than either the PAL16L8 or the PAL16R4. Since pins 14–23 can be used as inputs, combinational outputs, or registered outputs, the following inequalities describe the possibilities.

If the chip is used to implement a CL function, the constraints are

$$n \leq 22, \quad m \leq 10$$

$$(n + m) \leq 22$$

TABLE 9.6 Applications for Combinations of Switch Settings for the PAL22V10 Chip

Name	Output Connection	Feedback Connection	Application
INP	None	Pin	Use pin as input only
COCF	Combinational	Combinational	Combinational output and/or combinational feedback
COIF	Combinational	Pin (input)	Bidirectional pin implementing a combinational output
RORF	Register	Register	Typical state machine controller
ROIF	Register	Pin (input)	Bidirectional pin with registered output. Bus applications

For sequential circuits, the constraints are

$$n \leq 21, \quad m \leq 10, \quad k \leq 10$$

$$(m + k) \leq 10, \quad (n + m + k) \leq 21$$

because the clock signal for the flip-flops uses one of the input pins (pin 1).

Table 9.7 shows representative sizes for circuits that will fit onto the PAL22V10 chip.

TABLE 9.7 Representative Circuit Sizes That Will Fit onto a PAL22V20 Chip

Representative CL Circuits

M	n	Number of Product Terms per Output	Delay
1	21	16	T_A
1	20	31	$2T_A$
1	12	111	$2T_A$
2	20	16–16	$2T_A$
2	12	16–96, 29–83, 40–72, 49–63, 56–56	$2T_A$
3	19	14–16–16	T_A
3	12	16–16–81	$2T_A$
3	12	37–38–38	$2T_A$
5	12	15–19–23–27–31	$2T_A$
5	17	12–14–14–16–16	T_A
10	12	8–10–12–14–16–16–14–12–10–8	T_A

Representative Sequential Circuits

M	K	n	Number of Product Terms per (Output)–(State Variable)	
1	3	17	(16)−(14–16–16)	T_A
1	3	13	(25)−(25–25–25)	$2T_A$
1	3	11	(88)−(8–8–10)	$2T_A$
1	3	11	(31)−(27–28–28)	$2T_A$
2	3	16	(16–16)−(12–14–14)	T_A
2	3	13	(16–16)−(19–23–23)	$2T_A$
2	3	11	(44–45)−(8–8–10)	$2T_A$
2	3	11	(23–23)−(23–23–23)	$2T_A$
4	4	13	(14–14–16–16)−(10–10–12–12)	T_A
4	4	11	(12–12–29–29)−(8–8–10–10)	$2T_A$
5	5	11	(12–14–14–16–16)−(8–8–10–10–12)	T_A
2	8	11	(16–16)−(8–8–10–10–12–12–14–14)	T_A

9.1.10 FPGA Technologies

Due to the relatively high complexity of FPGAs, almost all FPGAs use CMOS process technology because of its high density and low power characteristics. Currently, there are two popular FPGA programming technologies, static RAM (SRAM), and anti-fuse.

The anti-fuse device gets its name from the fact that its electrical properties are the dual of the electrical properties of a fuse. The anti-fuse is a pair of conducting plates separated by a dielectric insulator, similar to a small capacitor. By contrast, the fuse is a pair of terminals separated by a thin conducting wire. A fuse is programmed by passing a high current through the thin wire causing the wire to heat up and melt, producing an open circuit where a short circuit previously existed. Fusible link technology is used in many PLA and EPROM devices. By contrast, the anti-fuse is programmed by applying a high voltage across the dielectric insulator that permanently breaks down the dielectric insulator, producing a short circuit where an open-circuit previously existed. Both fusible-link and anti-fuse devices are nonvolatile which makes them particularly well-suited for use in extreme environments, such as space and other high radiation environments. Anti-fuse technology also provides higher speed operation than other technologies.

SRAM chips are volatile (i.e., they lose their program when power is removed) and have lower density and slower speed than anti-fuse chips. On the positive side, SRAM chips are lower cost, re-programmable and, therefore, dynamically reconfigurable. In the current marketplace, SRAM chips have captured most of the popular commercial market.

9.1.11 FPGA Architectures

Figure 9.15 shows a high-level layout of an FPGA chip. Each chip contains a two-dimensional array of identical configurable logic blocks (CLBs). The FPGA in Figure 9.15 has 64 CLBs arranged in an 8×8 array. The user can program the CLBs to implement specific combinational or sequential logic functions. A programmable interconnect structure is permanently placed in the space between CLBs. The user programs switches that make desired connections between his programmed CLBs either by setting SRAM bits or by permanently closing the anti-fuses. Programmable I/O blocks are located around the perimeter of the chip that allows the user to connect signals to pins on the chip.

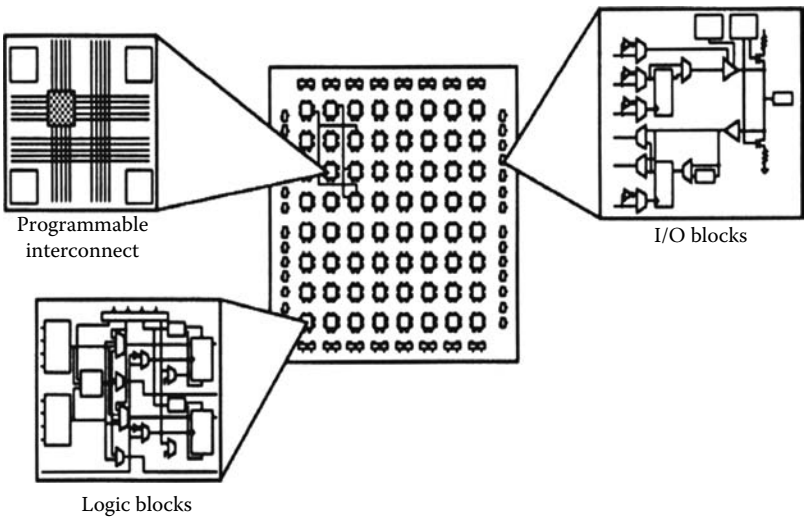


FIGURE 9.15 FPGA architecture. (Courtesy of Xilinx, San Jose.)

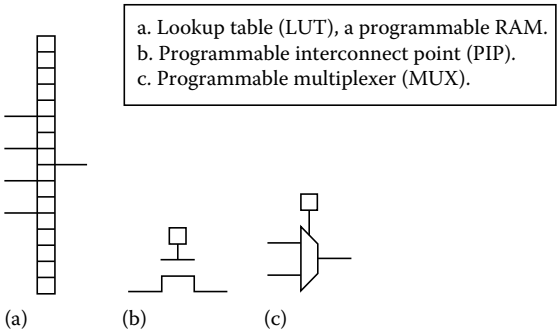


FIGURE 9.16 Programmable FPGA elements. (Courtesy of Xilinx, San Jose.)

For convenience, in this paragraph we describe FPGA elements using SRAM terminology. Anti-fuse devices have similar components. FPGAs have three basic programmable elements, illustrated in Figure 9.16. The lookup table (LUT) is a programmable RAM. The LUT shown in Figure 9.16a is a 16×1 RAM. It has four address inputs and one data output. It is programmed by storing either a logic 1 or a logic 0 in each RAM location. The value at a specific location is read out (looked up) by applying the address at the RAM inputs. A programmable interconnect point (PIP) is simply a CMOS pass transistor with a programmable SRAM bit controlling the gate signal. If a connection between the two transistor terminals for the PIP in Figure 9.16b is desired, a logic 1 is stored in the SRAM control bit. If no connection is desired, a logic 0 is stored in the SRAM control bit. SRAM bits also control the address lines of programmable MUXs. The MUX in Figure 9.16c has two input lines and therefore can be controlled by one SRAM bit. Programming the SRAM control bit to be logic 0 connects the upper MUX input to the MUX output. Programming the SRAM control bit to be logic 1 connects the lower MUX input to the MUX output.

To illustrate the principles, consider the minimal CLB shown in Figure 9.17. This CLB has three input signals (A, B, and C) and one output signal (X). The CLB has one 8×1 LUT and one D flip-flop with a reset input (R). It has three 2×1 programmable MUXs with SRAM control bits labeled M1, M2, and M3, respectively. The MUX controlled by M3 connects either the LUT output (F) or the bit stored in the flip-flop (Q) to the CLB output (X). If $M3 = 0$, the D flip-flop is bypassed and the CLB will implement the CL function stored in the LUT. If $M3 = 1$, the CLB will implement a sequential function. The MUX controlled by SRAM bit M2 selects either the LUT output (F) or the input signal (C) as the reset signal for the flip-flop. The MUX controlled by SRAM bit M1 selects either input signal C or the bit stored in the flip-flop (Q) as the third address input (E) to the LUT. Inputs *A* and *B* are permanently connected to two of the LUT address lines.

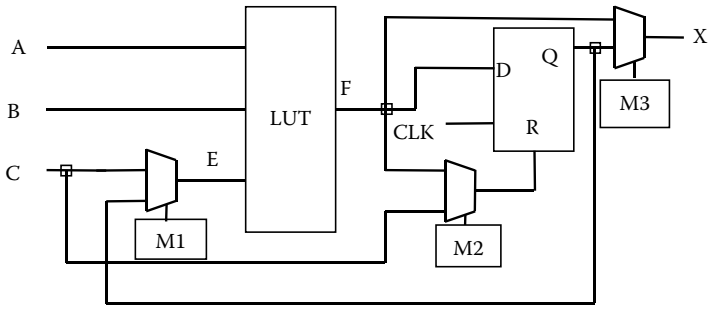


FIGURE 9.17 A minimal CLB.

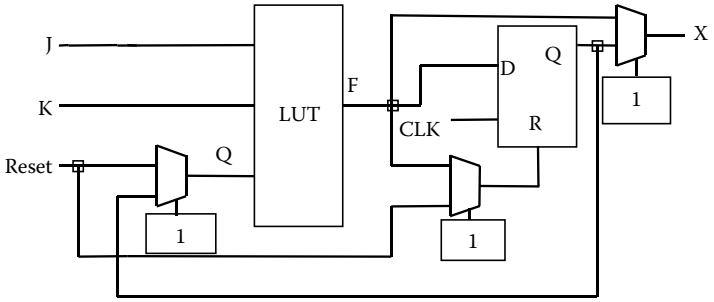


FIGURE 9.18 Minimal CLB programmed to be a JK flip-flop.

TABLE 9.8 Contents of LUT for Programmed CLB

J	K	Q	F
0	0	0	0
0	0	1	1
0	1	0	0
0	1	1	0
1	0	0	1
1	0	1	1
1	1	0	1
1	1	1	0

To illustrate how the minimal CLB in Figure 9.17 could be programmed, we will show how to program it to implement a JK flip-flop. M3 will be 1 to select the flip-flop output as the CLB output. M2 will be 1 to select input C as the flip-flop reset signal. M1 will be 1 to select the flip-flop output (Q) as the E input to the LUT. Input A will be designated as the J input and input B will be designated as the K input for the flip-flop. Figure 9.18 shows the programmed CLB with all signal names related to their use in the JK flip-flop. Table 9.8 shows how the LUT must be programmed to implement the function of a JK flip-flop.

Figure 9.19 shows an actual CLB in the XC4010XL chip, a small FPGA chip manufactured by Xilinx. This CLB uses the same small set of elements that we used in the minimal CLB. This CLB contains 2 D-flip-flops, 3 LUTs, and 16 programmable MUXs.

If the user had to directly program each CLB and each PIP in the interconnect structure, the task would be formidable. However, most chip manufacturers provide software packages that allow the user to specify the device function using a variety of high-level abstractions. In the next section, we will discuss this process in more detail.

9.1.12 Design Process

From the previous discussion, it is clear that fitting designs to PLD chips is a complex process. PLD manufacturers and third-party vendors market software packages that help engineers map designs onto chips. Selecting a package appropriate for a particular design environment is a critical decision that will significantly affect the productivity of the design group.

There are basically three types of development system packages: user designed packages, vendor designed packages, and universal packages. Since these programs are very complex and require many

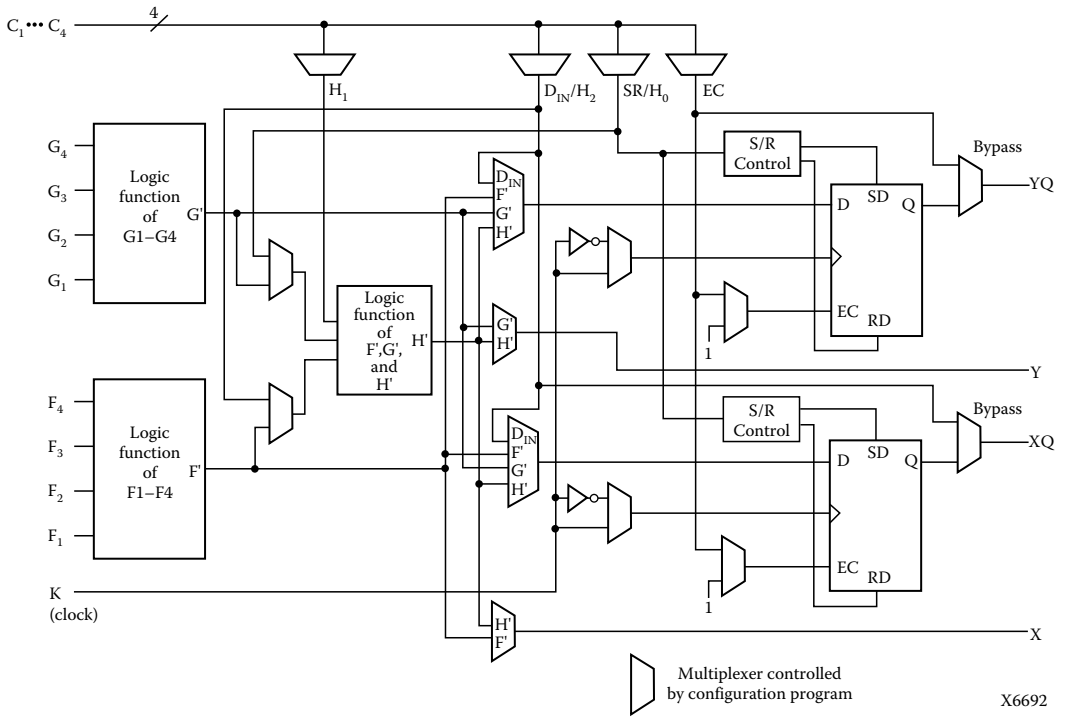


FIGURE 9.19 Actual CLB for the XC4010XL chip. (Courtesy of Xilinx, San Jose.)

years of effort to develop, most design groups will lack the time and resources to develop their own. Many vendors provide design aids that are specific to a particular product line. There is a great danger in becoming dependent upon one vendor's products because new products in this field appear frequently. Clearly, a universal design package that supports a wide variety of product lines is most desirable. A variety of development systems with different features, capabilities, and price is available.

Figure 9.20 shows a flow diagram for the process typically used to design PLAs and FPGAs. ''Design entry'' refers to the use of an editor to create a source file that specifies the functional behavior of the device. High-level simulation verifies correct functional behavior of the device. ''Logic synthesis'' refers to the process of implementing the design using the primitive elements present on a specific chip, such as gates, flip-flops, registers, etc. Most development systems support prelayout simulation at this level to verify that the design still functions correctly. ''System partitioning'' and ''mapping'' refers to the process of grouping blocks of primitive elements into sets that map directly into major chip structures, such as CLBs in FPGAs or AND–OR arrays in PLDs. ''Place and route'' refers to mapping the structures into specific locations on the chip and making connections between them. The software package then performs a timing analysis on the final design to verify that design timing specifications are met. Finally, the chip is configured by generating an output file that can be read by the chip programmer.

9.1.12.1 Design Entry

It is essential for a universal development system to have a variety of design entry modes. Many vendors market the more complex design entry modes as optional features. This section describes some of the more common design entry modes and their value to PLD and FPGA designers.

"Boolean equations" are an important method of design entry. A PLD design system must support Boolean equation entry because the AND–OR arrays on PLD chips directly implement Boolean equations. Almost all PLD designers will use Boolean equation entry extensively. Boolean equation entry is also useful for FPGA designs.

"Truth table" entry allows specification of a CL function by defining the output for each of the 2^n input combinations. This form is particularly valuable if "don't care" entries exist. Truth table entry is most commonly used for functions with a small number of input variables that are not easily described by Boolean equations. Code converters, decoders, and LUTs are examples. A good design tool will support truth table entry.

"Symbolic state machine" entry is crucial for PLD and FPGA designers because both PLDs and FPGAs are often used to implement state machine controllers. Current tools have features described as state machine entry that vary dramatically in form and usefulness. Before selecting a tool, the specifics of the state machine entry format should be carefully investigated. The most useful formats allow symbolic representation of states and specification of state transitions using some form of conditional statement such as "if_then_else," or "case." Relational operators are also useful in this context. The tool should perform automated state assignment and should fit the state variable equations to the target chip.

"State diagrams" using graphics are useful, but not essential. This feature is mainly a convenience, provided that symbolic state machine entry is available.

"Schematic" entry is a widely accepted way to describe logic systems. To be useful, it must be combined with a powerful partitioning and mapping tool that can fit the circuit onto chips. Schematic entry is useful to convert existing gate level designs into PLD or FPGA implementations.

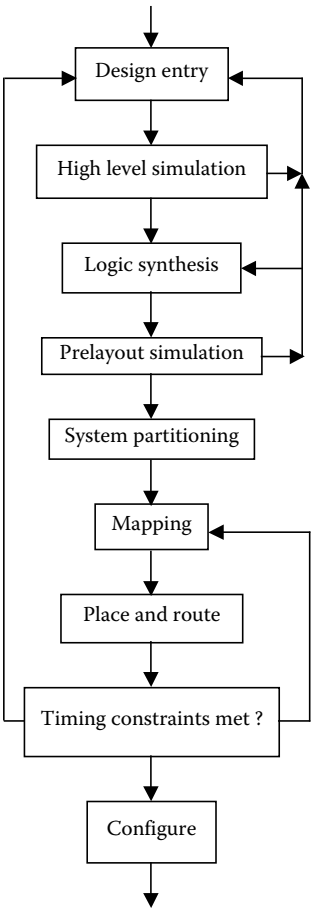


FIGURE 9.20 Design process.

"Hardware description language (HDL)" entry is potentially the most useful of all methods. Popular HDL languages are VHDL, Verilog, and System C. Using these languages, a designer can specify an executable specification of his device. Mature simulators exist for all of these languages that allow functional verification of the high level HDL design.

9.1.12.2 Logic Synthesis

Logic synthesis is the process of transforming a given description of a device produced by one of the design entry methods described in the previous section into an equivalent netlist using primitive components. For example, the process of transforming a symbolic state machine description or an HDL description into a netlist is an example of logic synthesis. The power of the synthesis algorithms in a development system is perhaps the most important feature of the system.

HDL synthesis tools are beginning to be mature enough for use in both PLD and FPGA designs. Very good synthesis tools exist for subsets of the popular HDL languages. Full synthesis tools for all language constructs are still in the research phase. In the next section, we will illustrate how to synthesize FPGA designs using the VHDL language.

"Logic minimization" is obviously an essential process in a PLD development system because the number of product terms per output gate on PAL chips is limited. Recall that the goal of logic minimization for PLD designs is to reduce the number of product terms, not the size of the product terms. Classical logic minimization algorithms use cost functions that reward reduction of the number of gate inputs. This is important for TTL gate implementations, for example, because an eight-input gate costs

about four times as much as a four-input gate. In PLD designs, the number of gate inputs does not matter. Each product term consumes one product line in the chip. A one-literal product term, such as X, costs exactly the same as a ten-literal product term, such as ABCDEFGHIJ. Therefore, traditional logic minimization programs, such as Espresso, need to be modified for PLD development. If product terms can be shared among different outputs, then multiple output minimization is necessary. However, for most PAL devices, the product terms cannot be shared; therefore, single output minimization algorithms are sufficient. Single output minimization algorithms are much less complex and take much less time to execute than multiple output minimization algorithms. Therefore, systems that do single output minimization result in higher productivity. Therefore, be careful of systems that advertise well-known traditional logic minimization algorithms to market their products, especially if multiple output minimization is stressed.

''Equation factoring,'' which is sometimes called ''multiple level minimization,'' is essential in order to fit large functions onto PLD chips using multiple cells combined with feedback lines inside the chips. This feature is missing from most vendor PLD development systems. However, in order to provide effective automated PLD synthesis, this operation is absolutely necessary. In most current PLD development systems, the designer must interact with the synthesis program to implement multiple level minimization. Such interaction requires extensive skill from the user of the software package.

9.1.12.3 Simulation of Designs

All good development systems include some form of simulation capability. The simulators vary widely in scope, user interface, and general usefulness.

Behavioral simulation allows high-level design descriptions to be simulated independent of implementation. Behavioral simulators verify the input–output behavior of the device. Correct behavioral simulation verifies the correctness of the algorithms prior to mapping the design to specific hardware components.

Device simulators verify the function of the design after mapping the design to a specific chip but before actually programming the chip. This is the most common type of simulator in current PLD development systems. A device simulator will construct a software model of the target PLD architecture, map the design to that architecture, and then simulate the behavior of the specific PLD. The better simulators will provide timing information as well as functional information.

9.1.12.4 Mapping Designs to Chips

System partitioning, mapping, place and route, and configure functions are usually performed by vendor-specific development software. These software packages usually accept a netlist as input and produce an output file that can be read by the programming tool.

```
SynchronousRegProcess: process (CLK)
 begin
  -- No other statements here
  if (CLK'event and CLK='1') then
   if RESET ='0' then
    AREG <='0';
   else
    AREG <= A;
   end if;
  end if;
  -- No other statements here
 end process;
```

FIGURE 9.21 VHDL code for register with synchronous reset.

9.1.13 VHDL Synthesis Style for FPGAs

Since HDL synthesis is one of the most popular ways to design FPGAs, we will show representative synthesis techniques for VHDL, one of the common HDL languages. The user may use these examples as templates to write code that will synthesize successfully.

9.1.13.1 Registers and Flip-Flops

Figure 9.21 shows a VHDL template for a register, AREG, with synchronous reset signal (RESET) and data input, A. In VHDL, entities are called processes. This code defines a process named

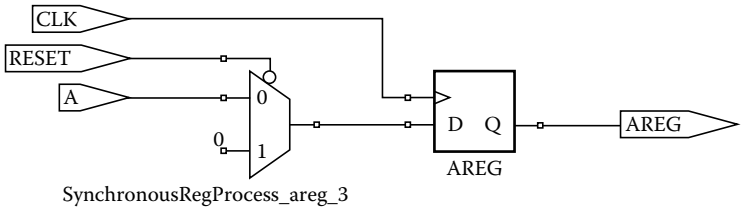


FIGURE 9.22 Synthesis of VHDL code in Figure 9.21.

SynchronousRegProcess. The list of signal names in the parentheses following the word process is called the sensitivity list for the process. The sensitivity list for a register process must contain the clock signal for the process. The notation, CLK'event, is a Boolean expression that is TRUE when signal CLK changes value. This change may either be a rising edge or a falling edge. Combining CLK'event with CLK = '1' specifies a rising edge triggered clock. The notation *AREG <= A* specifies that the current value of input signal A is assigned to register AREG on the rising edge of signal CLK. Similarly, the notation, AREG <= '0', implies that signal AREG is reset to 0 on the rising edge of CLK when RESET = '0'. Current synthesis semantics dictate the following constraints on the VHDL code.

1. The "if" statement that identifies the clock signal must be the only top level statement in the process. No other statements may come before or after this "if" statement, as indicated by comments in the VHDL code.
2. The condition that identifies the clock signal for the register (CLK'event and CLK = '1'), must contain no other conditions. For example, we could not implement the synchronous reset by writing (if CLK'event and CLK = '1' and RESET = '0' then...).
3. Only one clock edge may occur in each process. That is, it is unacceptable to have both CLK1'event ad CLK2'event in the same process.

Figure 9.22 shows how a synthesis tool will synthesize this VHDL code into a clocked register. Note that the RESET signal is low active (a logic 0 causes a reset).

Figure 9.23 shows a VHDL template for a register, AREG, with an asynchronous reset. The same restrictions that were listed for the synchronous register also apply here. Notice that the primary differences are that, in the asynchronous register, signal RESET must be in the sensitivity list and the RESET test occurs before the test for the clock edge, whereas the test for RESET in the synchronous register occurs after the test for the clock edge. Figure 9.24 shows how a synthesis tool synthesizes the VHDL code in Figure 9.23 into a register with an asynchronous reset signal.

```
Asynchronous_Reset: process (CLK, RESET)
 begin
  if RESET = '0' then
   AREG <= '0';
  elsif (CLK'event and CLK = '1') then
   AREG <= A;
  end if;
 end process;
```

FIGURE 9.23 VHDL code for a register with asynchronous reset.

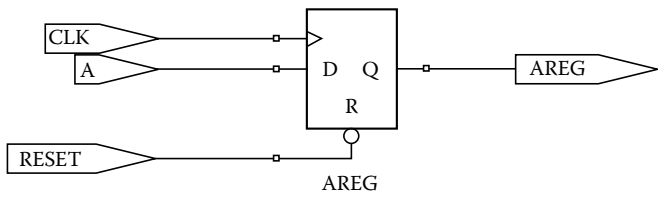


FIGURE 9.24 Synthesis of VHDL code in Figure 9.23.

TABLE 9.9 Ones Counter Truth Table

A(2)	A(1)	A(0)	C(1)	C(0)
0	0	0	0	0
0	0	1	0	1
0	1	0	0	1
0	1	1	1	0
1	0	0	0	1
1	0	1	1	0
1	1	0	1	0
1	1	1	1	1

```
process(A)
begin
  case A is
    when "000" => C<= "00";
    when "001"|"010"|"100" => C<= "01";
    when "011"|"101"|"110" => C<= "10";
    when "111" => C<= "11";
    when others => null;
  end case;
end process;
```

FIGURE 9.25 VHDL code for combinational logic.

9.1.13.2 Combinational Logic

Consider the ones counter truth table of Table 9.9. The input is a 3-bit vector, A, and the output is a 2-bit vector, C. The output reflects the number of ones in the input vector. For example, if A = 111, then C = 11 indicating that A has 3 ones.

There are many ways to represent CL using VHDL. The most direct way is to use a case statement, as shown in Figure 9.25. This type of code is directly related to the truth table format. It simply specifies the output for each combination of the inputs. Figure 9.26 shows how a synthesis tool synthesizes the code in Figure 9.25. Note that it takes the unusual approach of using a MUX as part of the circuit. The reason for this is that the target FPGA chip has many MUXs in its CLBs (see Figure 9.19).

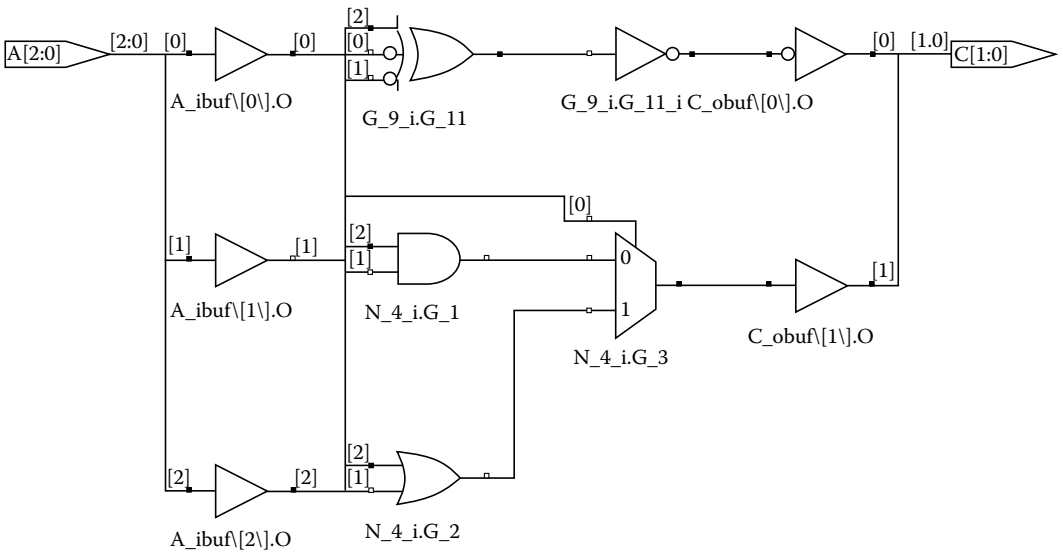


FIGURE 9.26 Synthesis of VHDL code Figure 9.25.

9.1.13.3 Latches

Figure 9.28 shows that the VHDL code in Figure 9.27 synthesizes as a latch with output signal name Q. Note that the process sensitivity list must include both the latch input data signal name, DATA, and the latch enable signal name, ENABLE. Also, the code for a latch may not contain an expression referring to an edge, such as ENABLE'event. The reason that the synthesis tool produces a latch is that a new value is not assigned to signal Q during every call to the process. A new value is only assigned when ENABLE = '1'. Otherwise, the signal Q must not change, i.e., it must retain its old value when ENABLE = '0'. This action requires a latch.

Figure 9.30 shows that the VHDL code in Figure 9.29 synthesizes as CL and does not produce a latch. In this VHDL code, signal Q is assigned a new value every time the process is called. Therefore, the synthesized circuit is CL instead of a latch.

9.1.14 Synthesis of State Machines

Figure 9.31 shows a VHDL template for a state machine. First, the code includes a declaration of a data type called STATE_TYPE. This data type is simply a list of the names of the states. The names should be chosen to reflect the purpose of the state, such as INITIAL, IDLE, TRANSMIT, RECEIVE, etc. Next, a signal, STATE, is declared to be of type STATE_TYPE. Signal STATE keeps track of the current state of

```
LATCH: process (ENABLE, DATA)
  begin
    if (ENABLE = '1') then
      Q <= DATA;
    end if;
  end process;
```

FIGURE 9.27 VHDL code for a latch.

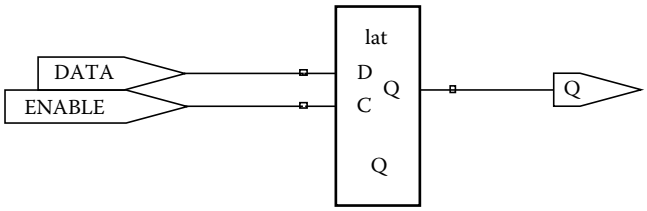


FIGURE 9.28 Synthesis of VHDL code in Figure 9.27.

```
CLP: process (ENABLE, DATA)
  begin
    if ENABLE = '1' then
      Q <= DATA;
    else
      Q <= '0';
    end if;
  end process;
```

FIGURE 9.29 VHDL code for combinational logic.

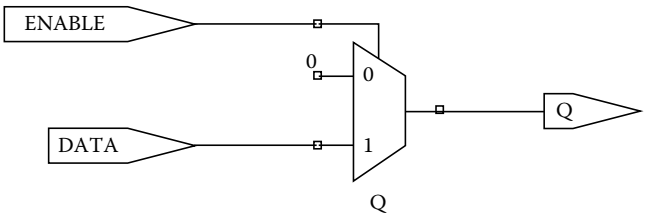


FIGURE 9.30 Synthesis of VHDL code in Figure 9.29.

```
architecture FSM of NAME is
    type STATE_TYPE is (S0, S1,..., Sn);
    signal STATE: STATE_TYPE ;
        -- other_signal_declarations
begin
STATE_PROCESS: process (CLK, RESET)
  begin
    if RESET='0' then
      STATE <= S0; -- The initial state
      -- Insert Reset statements
    elsif CLK'event and CLK = '1' then
      case STATE is
        when S0 = >
             -- Data_Section
             -- Control_Section
        when S1 =>
             -- Data Section
             -- Control Section
        when others =>
             -- Actions
      end case;
    end if;
  end process;
OUTPUT_PROCESS: process (STATE) begin
    case STATE is
      when S0 = >
           -- Output_Signal_Assignments
      when S1 =>
           -- Signal Assignments
      when others =>
           -- Signal Assignments
    end case;
  end process;
end FSM;
```

FIGURE 9.31 VHDL template for a state machine.

the state machine. The architecture consists of two processes. Process STATE_PROCESS updates the current state on each positive transition of clock signal CLK. Process OUTPUT_PROCESS updates the state machine outputs whenever there is a change in state. As written, these processes implement a Moore state machine. To design a Mealy state machine, simply add the machine input signals to the OUTPUT_PROCESS sensitivity list. The state machine has a low-active asynchronous RESET signal that initializes the state machine to state S0. A case statement performs data transfer statements and computes the next state based on the current state. For other approaches to using high-level languages to design digital systems, see Ref. [8].

References

1. *Programmable Logic Data Book*, Texas Instruments, Dallas, TX.
2. *Programmable Logic*, Intel Corporation, Mt. Prospect, IL.
3. C. Alford, *Programmable Logic Designer's Guide*, Howard W. Sams & Company, Indianapolis, IN, 1989.
4. H. Katz, *Contemporary Logic Design*, Benjamin/Cummings Publishing Company, Redwood City, CA, 1994.
5. L. Pappas, *Digital Design*, West Publishing Company, St. Paul, MN, 1994.
6. D. Pellerin and M. Holley, *Practical Design Using Programmable Logic*, Prentice Hall, Englewood Cliffs, NJ, 1991.
7. J. F. Wakerly, *Digital Design, Principles & Practices*, 2nd ed., Prentice Hall, Englewood Cliffs, NJ, 1994.
8. J. R. Armstrong and F. G. Gray, *VHDL Design: Representation and Synthesis*, 2nd edn., Prentice Hall, Englewood Cliffs, NJ, 2000.
9. V. P. Nelson, H. T. Nagle, B. D. Carroll, and J. D. Irwin, *Digital Logic Circuit Analysis and Design*, Prentice Hall, Englewood Cliffs, NJ, 1995.

9.2 Clocking Schemes

Wayne D. Grover

9.2.1 Introduction

Advances in very large-scale integrated (VLSI) processing technology, particularly CMOS, have resulted in nanometer-scale processes with applications at clock speeds of several gigahertz. For example, at the time of this revision the state of the art is fairly well represented by the Intel Core2 Duo processor chip, which is implemented in 65 nm CMOS, consists of 376 million transistors, and clocks at 2.66 GHz. New design challenges must be mastered to realize systems at an ever-increasing clocking rates and circuit sizes. In particular, clocking-related issues of skew, delay, power dissipation, and switching noise can be design-limiting factors. In large synchronous designs, the clock net is typically the largest contributor to on-chip power dissipation and electrical noise generation, particularly ''ground bounce,'' which reduces noise margin. Ground bounce is a rise in ground potential due to surges of current returning through a nonzero (typically inductive) ground path impedance. At the board and shelf level, clock distribution networks can also be a source of electromagnetic emissions, and may require considerable delay tuning for optimization of the clock distribution network.

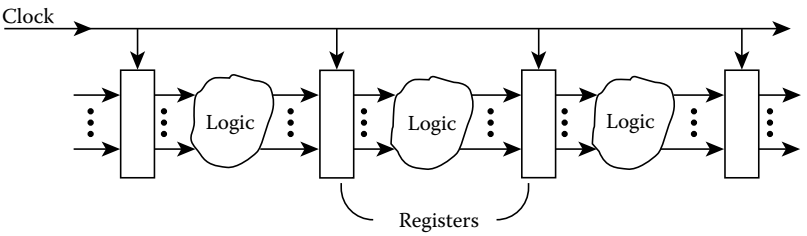
In the past, multiphase clocking schemes and dynamic logic structures helped minimize transistor count, but this is now less important than achieving low skew, enhancing routability, controlling clock-related switching noise, and providing effective CAD tools for clock net synthesis and documentation. For these reasons, a shift has occurred toward single-phase clocking and fully static logic in all but the largest custom designs today. In addition, phase-feedback control schemes using phase-locked loops (PLLs) are becoming common, as are algorithmic clock-tree synthesis methods.

This section focuses on the issues and alternatives for on-chip and multichip clocking, with the primary emphasis on CMOS technology. We first review the fundamental nature and sources of skew and the requirements for the clocking of storage elements. We then outline and compare a number of ''open-loop'' clock distribution approaches, such as the single-buffer, clock trunk, clock ring, H-tree, and balanced clock tree approaches. PLL synchronization methods and PLL-based clock generation are then outlined. In closing, we look at future technologies and developments for high-speed clocking. The concepts and methods of this section apply to many circuit technologies on- and off-chip. However, we emphasize CMOS because CMOS processes (including bi-CMOS) presently represent the vast majority of digital VLSI designs and are expected to continue to do so.

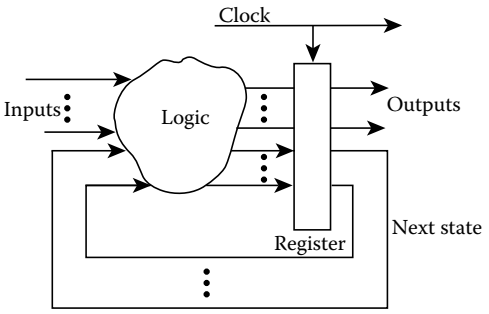
Asynchronous, self-timed, and wavefront array systems are outside the scope of this chapter. These approaches aim to minimize the need for low-skew synchronous clocking. However, truly asynchronous modules tend to require a large overhead in logic for interaction with each other, so that speed, size, and power often suffer relative to synchronous design. Nonetheless, self-timing can be an effective approach for random-access memory (RAM and ROM) cells, to which considerable optimization effort can be invested for reuse in many designs. Self-timed methods should be considered the alternative to fully synchronous design in large, highly modularized systems, particularly where well-defined autonomous modules have relatively infrequent interactions. The main issues in self-timed systems are the possibly high delay required to avoid metastability problems between self-timed modules, and the circuit costs of the synchronization protocol for intermodule communication.

9.2.2 Clocking Principles

Most of us accept the clocked nature of digital systems without question, but what, fundamentally, is the reason for clocking? Any digital system can be viewed either as a pipeline or as an FSM architecture, as outlined in Figure 9.32. In the pipelined architecture clocked sections are cascaded, each section comprising an asynchronous CL block followed by a latch or storage element that samples and holds the logic state at the clock instant. In the FSM, the only difference is that the next state input and the system outputs are determined by the asynchronous logic block, and the sampled next state value (S) are fed back into the CL. The FSM can therefore be conceptually unfolded and also represented in a pipeline fashion. The fundamental reason for clocking digital systems is seen in this pipelined abstraction of a digital system: it is to bring together and retain coordination among asynchronously evolved intermediate results. With physical delays that are temperature, process, and input dependent in CL, we need to create ''agreed-upon time instants'' at which all analog voltages in a system are valid when interpreted as Boolean logic states. Clocking deals with delay uncertainty in logic circuit paths by holding



(a) Pipelined architecture



(b) Finite-state machine atchitecture

FIGURE 9.32 Architecture of digital systems.

up the fast signals and waiting for the slower signals so that both are valid before they are again combined or interact with each other. Without this coordination purely asynchronous logic would develop severe propagation path differences, and be slow in repetitive operations. Ultimately, a valid state would evolve, but all inputs would have to be stable for the entire time required for this evolution. On the other hand, when the overall CL function is appropriately partitioned between clocked storage latches, system speed can approach the limit given by the delay of a single gate because each logic subblock is reused in each clock period.

From this, we obtain several insights: (1) only the storage elements of a digital system need become loads on the clock net (assuming state logic gates); (2) the system cannot be clocked faster than the rate set by the slowest combinational signal path delay between clocked storage elements; (3) any uncertainty in clock timing (skew) is indistinguishable from uncertainty in the settling time of the intervening CL; and (4) for a logic family to work, its storage elements must: (a) at no time be transparent (i.e., simultaneously connect input to output), (b) have a setup time less than $(T - t_{clk-Q})$ where T is the clock period and t_{clk-Q} is the clock-to-Q output delay of the same type of flop, and (c) have a hold time less than their clock-to-output delay. The last points may be better appreciated by considering that the CL block may be null, i.e., a zero-delay wire, such as in a shift-register.

An implication of [4(a)] is that two-phase nonoverlapping clocks, or an equivalent sequencing process, are fundamentals for the storage elements of a digital system. This may sound unusual to readers who have already designed entire systems with SSI and MSI parts, or in gate-array design systems, without having seen anything but single-phase edge-triggered flip-flops, latches, counters, etc. However, at least two clock phases (or clock-enabling phases) are internally required in any clocked storage device. An analogy is of a ship descending in elevation through a lock [27]. During the first clock phase, sluice gates ''charge'' the lock up to the incoming water level and open the input gate to bring a ship in. Throughout this phase, it is essential that the output gate is closed, or water will race destructively right though the lock. Only when the ship is entirely in the lock and the input gate is closed (isolating the input) can the output sluice gates be opened to equalize the water level to that on the outgoing side, allowing the ship to leave (producing a new output). Similarly, in a flip-flop or latch, the currently stored value, which appears at the output, must be isolated from the input while the input evolves to its next value.

9.2.2.1 Skew and Delay

Clock ''skew'' is defined, most generally, as the difference in time between the actual and the desired instant of active clock edge at a given clocked storage element. In the majority of designs in which the desired instant of clocking is the same at all storage elements, skew is the maximum difference in clock waveform timing at different latches. Clock skew is of concern because it ultimately leads to the violation of setup or hold times within latches, or to clock race problems in multiphase clocking. Furthermore, from a design viewpoint, uncertainty in clock timings must be treated as equivalent to actual clock skew. Skew or timing uncertainty are therefore equivalent to an increase in critical path logic delay. In either case, the clock period must be extended to ensure valid logic levels and proper setup/hold requirements relative to the clock time.

To illustrate the equivalence of skew (either actual of design timing uncertainty) to a loss of system speed, consider a 200 MHz process used in a design which has 25% skew (i.e., actual clock edge dispersion or, equivalently, uncertainty in clock timing is 1.25 ns). If a competitor uses the same process at 200 MHz and achieves 5% skew (0.25 ns), then the latter design has 20% more of each 5 ns clock cycle for settling CL paths. Alternatively, for the same logic functions, the low-skew system could be clocked at 250 MHz with the same timing margins as the high-skew system. Skew, therefore, represents a loss of performance relative to basic process capabilities developed at the great expense. However, skew reduction costs relatively little, and is in the logic designer's control, not the process developer's and yet it is directly equivalent to a basic enhancement in process speed.

Skew is usually of primary concern on-chip or within any module that is designed on the presumption of a uniform clock phase throughout the module. Clock ''delay,'' on the other hand, is the difference

between the nominal clock edge time at an internal flip-flop and the system clock, or timing reference, external to the chip. While skew is typically managed internal to the die, delay is of concern at the system level to ensure external setup and hold time requirements. Skew and delay may be independent in any given clock distribution scheme. For example, an on-chip clocking tree that yields essentially zero skew may, nonetheless, impart a high clock delay, which will be of importance at the system level. The ''early clock'' technique and some PLL methods (presented later) can be used to address on-chip clock delay problems.

9.2.2.2 Isochronic or Equipotential Regions

The clock distribution problem arises at all scales of system design, from on-chip clock distribution in VLSI and wafer-scale integration (WSI) to the synchronization of circuit packs tens of meters apart. These applications are unified as a generic problem: synchronous clocking of ''electrically large'' systems, i.e., systems in which propagation time across the system is significant relative to the clock period. In such systems:

$$D/v > k/f_{\text{app}} \tag{9.1}$$

where

D is the characteristic scale or distance of the system
v is the propagation velocity
f_{app} is the application clock frequency
k is the skew requirement as a fraction of the clock period

For all locations around a clock entry point at which Equation 9.1 is false, we can consider events to be essentially simultaneous (or, equivalently, the region is a single electrical node), and the clock can be distributed within such regions without delay equalization. The region over which clock can be distributed without any significant skew is also known as an equipotential [27] region, or an isochronic region [1]. In this section, we are concerned only with cases in which Equation 9.1 is true, but it is implicit that the clocked end nodes may be either clocked loads directly or a buffer that feeds a local isochronic region.

The diameter (and shape) of an isochronic region on-chips depends on the wire type employed for interconnection. To control skew on chip, we need to consider delay differences due both to wire lengths and to the lumped capacitive efforts of the driven loads. Where the RC time constant of the wiring interconnect τ_w is much less than the RC combination of the driving source resistance and the lumped capacitance of N clocked loads on the net ($\tau_{\text{net}} = R_s C_{\text{gate}} N = N\tau_g$), we can consider all points on a net to be isochronic, meaning that the net acts like one electrical node characterized by the total lumped capacitance of gates on the net. Wires on a chip are often modeled as distributed R_0C_0 sections, where R_0 and C_0 are the resistance and capacitance per unit length, respectively (see Figure 9.33). In such cases, the propagation delay for a wire of length l follows the diffusion equation [31].

$$\tau_w = R_0 C_0 l^2/2 \tag{9.2}$$

Therefore, if we consider a net of length l with N standard loads, we can consider the net to be isochronic if $\tau_w \ll N\tau_g$. From Equation 9.2, this implies

$$l \ll \sqrt{\frac{2NR_sC_{\text{gate}}}{R_0C_0}} \tag{9.3}$$

This relationship provides a guideline for the maximum length over which wire delays may be neglected relative to gate-charging delays. Based on typical values for a 1 μm process, $\tau_g < 500$ ps, isochronic regions for lightly loaded lines ($N = 1$) are up to 10,000λ for lines in third layer metal, 5,000 and 8,000λ

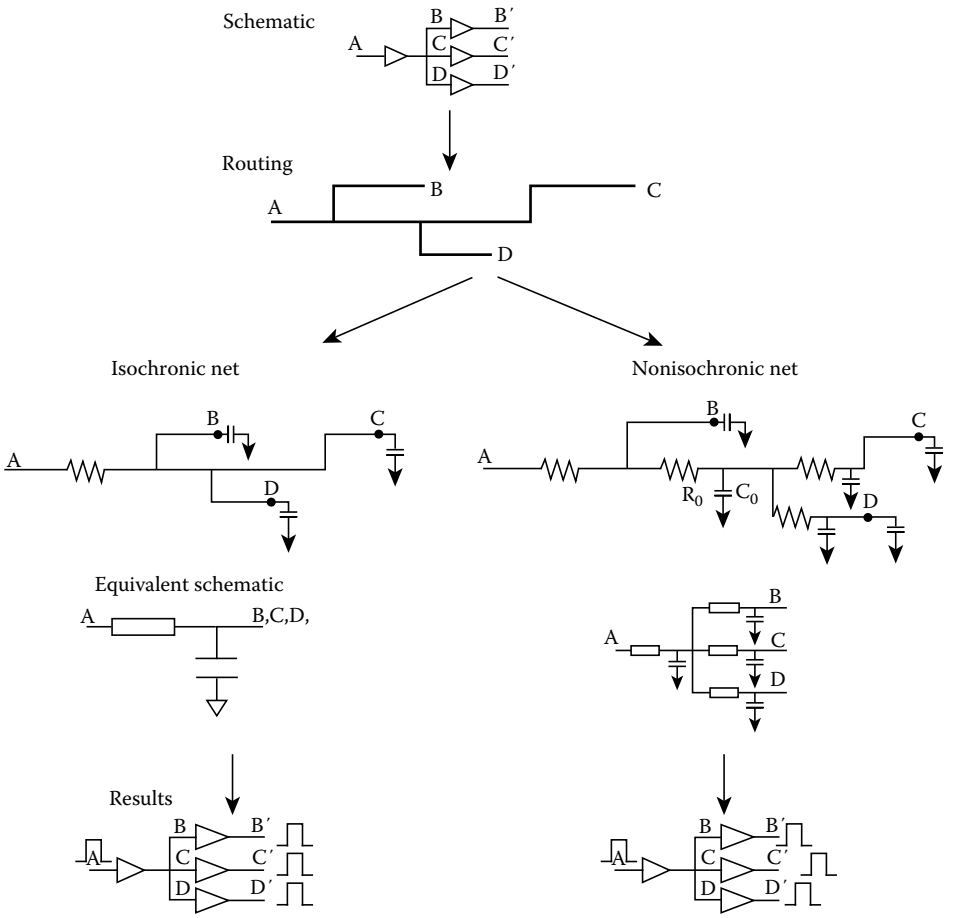


FIGURE 9.33 Isochronic and nonisochronic passive wiring nets.

for first and second layer metal, respectively, and 200λ for polysilicon wires, where λ is the minimum feature size of the process [31]. This illustrates the importance of distributing clock within metal layers to the greatest extent possible. Even a few short polysilicon links may introduce sufficient series resistance to drastically reduce the isochronic region for the given clock line. This also illustrates that if clock is distributed in metal layers, and is always buffered before exceeding the isochronic distance, it will be primarily differences in lumped capacitive loads and not wire length that determine clock signal delays, and hence relative clock skews.

9.2.2.3 Nature of Skew On-Chip

The concept of isochronic regions helps us understand the nature of clock skews in VLSI and helps explain why on-chip skews may be greater than those between off-chip points that are physically many more times distant. A key realization is that signals do not propagate at the "speed of light." If they did, absolute delays across even the largest chips (2 cm edges) would be subnanosecond and the isochronic diameter would easily encompass an entire die at clock speeds up to 200 MHz. Rather, on-chip propagation delay depends much more on the time needed for output drivers to charge the total lumped capacitance associated with all the gate inputs of the driven net. In other words, fanout and driver current

abilities have more to do with delay than path lengths. This is especially true when clock distribution is exclusively via metal layers, as is the norm in a modern design. On the other hand, off-chip, we route signals via impedance-controlled coaxial or microstrip lines, or via optical fiber, and these media typically do exhibit propagation velocities of 0.6–0.8 c. Therefore, off-chip, differences in physical propagation distances are the dominant source of skew, while on-chip, it is imbalances in driver loads that are the most common source of skew.

In on-chip cases in which wire diffusion delays and lumped capacitive effects are both significant, a difference in line length can also result in skew due to a different total wiring capacitance. In addition, equal length lines that go through different metallization layers or through polysilicon links will have different delays due to different levels of capacitive coupling to V_{ss} and different series resistances, especially in the case of polysilicon links. Accordingly, an important principal to simplify clock net design is to aim for buffering levels and fanouts that yield isochronic conditions for the passive wiring nets between buffered points on the clock net. This simplifies skew control in clock net design because attention then only need be paid to balancing loads on each buffer and to matching the number of buffers in each clock path. The alternative, in which passive wiring nets are not isochronic, requires detailed delay modeling of each wire path, taking into account the actual routing, the R_0C_0 of the wire type, the temperature, and the exact position of each lumped load on the wiring path. The important concept, however, is that by the choice of metal layers, line widths, and/or loadings, one can establish formally defined isochronic conditions on some or all portions of a complete clock net, which, in its entirety, is far too large to be isochronic. When fully isochronic subregions (such as a wide clock trunk) can be established, or even when a defined region is not isochronic but has delay that is simply and reliably predicted from position (such as on a clock ring), the remaining clock net layout and skew control problem is simplified, design risk is lowered, and pre- and postlayout simulations are more consistent because final routing of clock paths from these reference regions is shortened and overall uncertainty reduced. We shall see and use this principle in analyzing the popular clock distribution schemes that follow.

The skew that intrinsically arises from differences in RC time constants of either lines or gate loads is aggravated by threshold variations in buffers and clocked loads due to minute differences in electronic parameters and lithographic variation in line widths and lengths at different devices. Time-constant and threshold effects interact to give a worst-case skew, which is the difference between the time at which the voltage response of line with the slowest time constant, τ_{max}, crosses the threshold of the logic element with the highest threshold, V_{Tmax}, until switching of the device with the lowest threshold driven by the line with fastest RC time constant. Taking the difference of the earliest and latest switching times we have [32]

$$\delta = \tau_{min}\ \ln\left(1 - \frac{V_{Tmin}}{V_{DD}}\right) - \tau_{max}\ \ln\left(1 - \frac{V_{Tmin}}{V_{DD}}\right) \tag{9.4}$$

Equation 9.4 implies that a clock system design in which buffered electrical segments of the clock net have 10% variation in τ about τ_{nom}, and 10% variation of V_T about $V_{DD}/2$, will have an estimated skew of at least 17% of τ_{nom}.

9.2.2.4 Single-Phase Clocking

Clocks ultimately always drive a storage register or latch of some type. The form of clock signal(s) required in a system therefore depends on the type of latch or flip-flop element used and on properties of the CL circuits used. True single-phase clocking is the most complex clocking principles with which to design systems, and has traditionally not been used, although recent work has assessed some truly single phase logic families [2]. The reason for caution with single-phase clocking is that invalid states may be passed to the output in two ways, as shown in Figure 9.34a if the CL delay is less than T_h (i.e., too fast) or (Figure 9.34b) the CL delay is

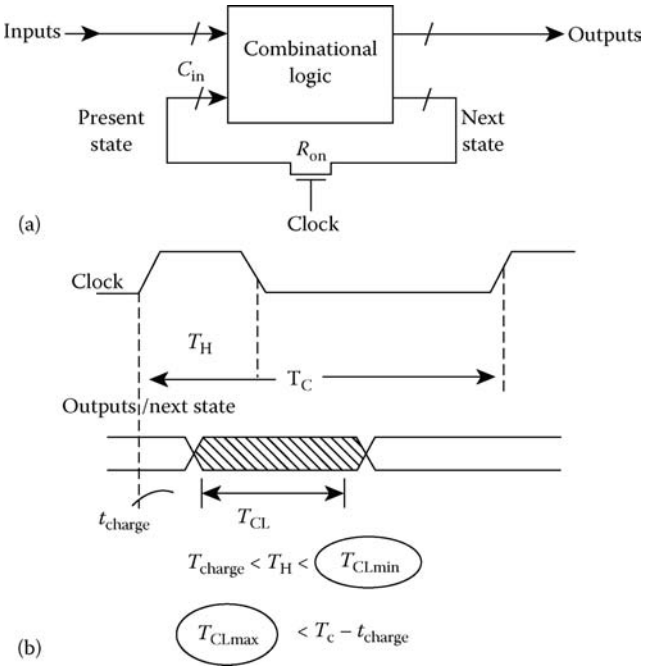


FIGURE 9.34 In single-phase clocking the CL path must be neither too slow nor too fast.

greater than $T_C - t_{charge}$ (i.e., too slow). In other words, a two-sided (min and max) constraint on logic path delay exists for single-phase clocking [27]. This means that although attractive to minimize total interconnect, buffer counts, and interphase skew is avoided, truly single-phase clocking involves a greater design risk and timing analysis complexity. Because of this, the most common overall clocking scheme is single-phase clock distribution with local generation of a two-phase clock.

9.2.2.5 Two-Phase Clocking

With two nonoverlapping clock phases, we can eliminate one of the risks of single-phase clocking, that of a logic path being too fast. On the first phase the input is made transparent and allowed to affect the CL, charging the inputs through R_{on} C_{in} in time t_{charge}. During this time, the CL outputs are isolated from the input latch. On the second phase, the new CL output values are stored by the second phase latch while the input latch is opaque, isolating inputs from the new values until the next phase one clock time. A nonoverlapping period between phases ensures that at no time does direct transparency occur from input to output. With two-phase nonoverlapping clocks, as shown in Figure 9.35a, we need to ensure only that the maximum delay in the CL is less than $T_C - t_{charge} - T_3 - t_{preset}$. It is essential that the nonoverlapping interval, T_2, be greater than zero, but T_3 can be arbitrarily short. When present, however, T_3 acts as an extra timing margin against skew.

It is obviously desirable to make T_2 as small as possible, but we can do so only when distributing two-phase clock directly if the interphase skew is less than T_2. In the worst case, the interphase skew may be twice the skew of each of the two clock phase nets individually. Skew, therefore, necessitates at least a 1:1 derating in speed for two-phase clocking in addition to the basic loss of clock cycle time for logic settling, to ensure correct operation of storage devices. If skews in the two clock phase nets are uncorrelated, however, the extra penalty could be as high as 2:1. Every nanosecond of skew in the clock net for each phase then not only reduces the basic critical path logic timing margin by 1 ns, but also adds 2 ns to the T_2 requirement.

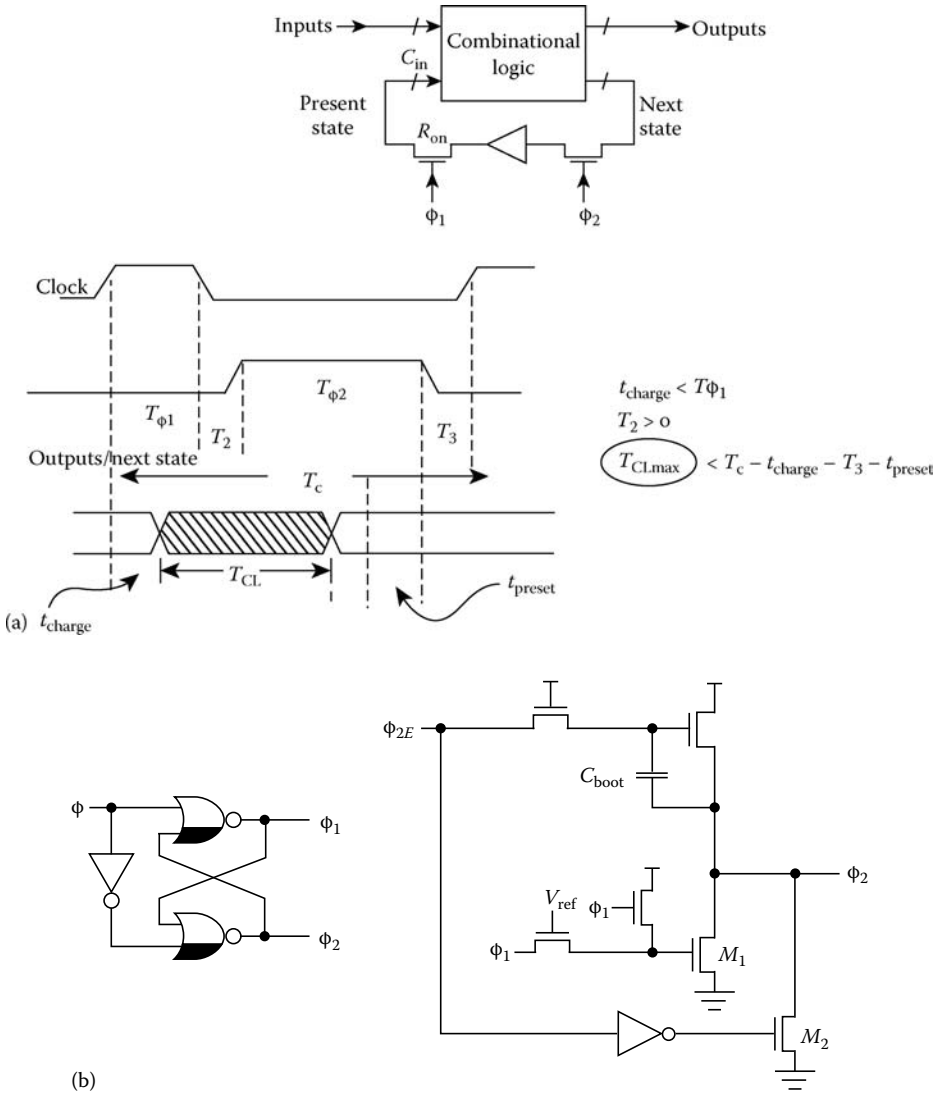


FIGURE 9.35 (a) With nonoverlapping two-phase clocks, no lower limit exists on the CL delay; and (b) generator for two-phase nonoverlapping clock and buffer circuit to ensure nonoverlap period. (From Glasser, L.A. and Dobberpuhl, D.W., *The Design and Analysis of VLSI Circuits*, Addison-Wesley, Reading, MA, 1985, 349.)

Therefore, in high-performance systems we have quite an incentive to distribute a single clock phase throughout the design and accept the extra logic required to generate two-phase clocks locally at each device (or small group of similar devices) that requires them.

9.2.2.6 Two-Phase Clock Generator Circuit

The canonical form of circuit to generate the local two-phase nonoverlapping clocks from a single phase clock is shown in Figure 9.35b. The feedback of ϕ_2 into NOR1 ensures that ϕ_2 must be low before ϕ_1 can go high after the single-phase ϕ_{in} input has gone low, and vice versa. A special clock buffer circuit is shown in Figure 9.35b, which helps ensure that a period of nonoverlap exists in the presence of the threshold variations in the driven loads [12]. It does this by using transistor M_1 to clamp the ϕ_2 output

low until far into the fall of ϕ_1. ϕ_2 is held low until ϕ_1 has fallen below ($V_{ref} - V_{thresh}$) to finally cut off M_1. V_{ref} might be set at 2 V in a 5 V process, thereby ensuring that ϕ_1 is well below the logic threshold of all clocked loads before ϕ_2 begins to rise, while at the same time minimizing but guaranteeing the existence of a nonoverlap period, which is lost processing time.

9.2.2.7 Multiple-Phase Overlapping Clocks

Generating and/or distributing nonoverlapping clocks with a minimal T_2 can be quite difficult in large systems. An alternative is to define three or more primary functional logic steps and use a similar number of overlapping clock phases. In this case, the multiple stages of clocking removes the need for the guaranteed nonoverlap period in two-phase clocking. Let us consider three-phase overlapping clocking. The principles generalize to any higher number of phases.

In three-phase clocking, the middle phase can be thought of as providing the nonoverlap time, which ensures time isolation between I/O activation for the module enabled on each clock phase. In fact, each phase plays a similar isolating role with respect to operations performed on its adjacent phases. Figure 9.36 illustrates the concept. The number of phases and the role for each phase typically reflects some natural cycle or step sequence of the basic system being designed; for example: bus input, add to accumulator, bus output.

In WSI systems, in which uncertainty in clock delays, circuit speeds, and interconnect impedances may be high [10], overlapping clock logic can give high tolerance to clock skew, and is compatible with self-timing in selected subcircuits. Three-phase overlapping clocking has a distinct advantage: no hazard exists unless all three clock phases overlap in time. In the presence of severe clock skew, this can be a major advantage. Although called overlapping clocks, the circuits still function if successive phases do not actually overlap, although speed is sacrificed if overlap is lost.

9.2.2.8 Overlapping Clock Phase Generator

Figure 9.37 illustrates a circuit for generating three-phase overlapping clocks. Phase overlap is ensured because it is the onset of each phase that kills its predecessor. A Johnston counter comprised of three static *D*-flip-flops generates the phase-enabling signals which sequence the actual generator stage, which is comprised of the three cross-coupled NOR gates. A deliberately limited positive-going drive ability of the enable input ensures that the Johnston counter exercises underlying rate and sequence control, while the output waveforms are determined by the interactions between the actual clock phase signals. While the enable inputs to each NOR are logically sufficient to drive the output high when the other input is low, they are arranged not to be able to drive the NOR output low on their own when the enable signal returns high. The output of phase *i* therefore stays high after its own enable signal has

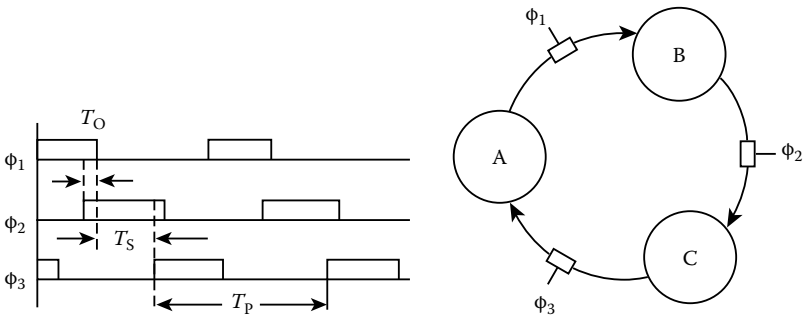


FIGURE 9.36 Principle of multiphase clocking. Outputs are isolated from inputs by other stages of nonactive logic even though any two active clock waveforms may overlap. (From Glasser, L.A. and Dobberpuhl, D.W., *The Design and Analysis of VLSI Circuits*, Addison-Wesley, Reading, MA, 1985, 352.)

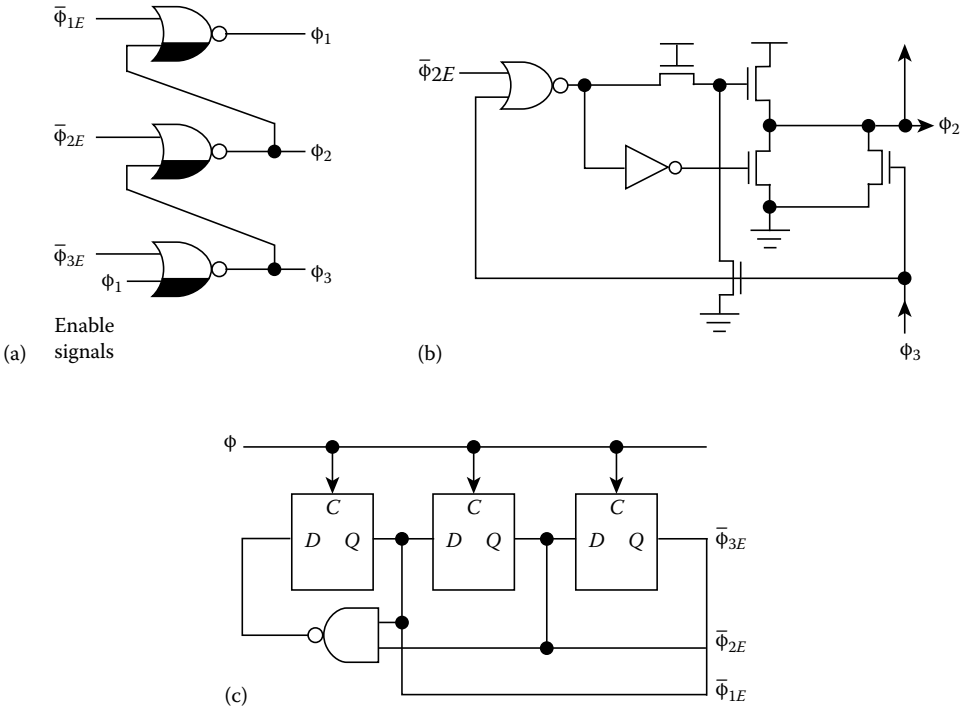


FIGURE 9.37 Three-phase clock generator logic and buffer design to ensure overlap. (From Glasser, L.A. and Dobberpuhl, D.W., *The Design and Analysis of VLSI Circuits*, Addison-Wesley, Reading, MA, 1985, pp. 348 and 354.)

disappeared (gone high) until the phase $i+1$ output is also high in response to the low-going phase $i+1$ enable. Figure 9.37 shows this logic and a NORing clock buffer circuit in which the phase $i+1$ signal is necessary to assist in returning the phase i output to zero.

9.2.2.9 Clocking Latches

A latch is a storage element which is level sensitive to the clock waveform. The latch output conforms to the input that is present while the clock waveform is at its active level for that latch, and then continues to hold that value when the clock level falls below the active level. Latches have setup and hold time requirements analogous to those in the flip-flops that follow. Circuit designs for high and low active latches are given in [31]. By combining latches of opposite active polarity, with logic between the stages, there can be two logic operations per clock period. In this type of operation, however, skew adds to the needed clock period as usual, but in addition any imbalance in the clock duty cycle requires a further margin because the minimum duty cycle half-width must remain greater than the worst-case logic delay. The clock edges also must be kept sharp enough that transparency never occurs between two successive latches working on opposite clock phases simultaneously, or that some minimum logic delay always exists between latches that exceeds the possible overlap time. The DEC ALPHA microprocessor was an example of a two-phase latch machine in which both phases drive latches that are active in the respective phases permitting logic evaluation twice per cycle. This is one case in which to control the very high transistor count, two-phase clock is distributed globally rather than generated at each module. The entire clock net on each phase is driven from a single large buffer placed at the center of the die where a PLL is also fabricated to advance the on-chip clocking phase relative to external bus timing, thereby compensating for buffer delay.

Two principles for maintaining 50% clock duty cycle in a latch machine are (1) whenever generating or phase-locking to a system clock, do so at twice the rate needed, then (frequency) divide by two. This results in 50/50 clock waveform, regardless of the original clock source waveform. (2) When repeatedly buffering clock in a chain, or when distributing clock through a hierarchy of clock buffers, use inverting clock buffers at each stage. Inverting the clock at every buffering stage inherently counteracts the effects of different rise and fall times in buffers. Otherwise, these can accumulate to extend or shorten the ON period of the clock waveform. For example, if a noninverting buffer has greater fall time than rise time, a clock path transiting several of these buffers will develop a broadened ON period. This effect is self-compensating in a chain of inverting buffers.

9.2.2.10 Clocking Flip-Flops

Flip-flops are more complex storage circuits than latches, but have no dependency on clock duty cycle because they are sensitive only during an active edge of the clock waveform. A rising edge *D*-flip-flop (for instance) updates its output to match its *D* input on the rising edge of the clock waveform. The *Q* output retains the updated value thereafter, regardless of further changes in the input or clock waveform (with the exception of another rising clock transition). A latch is a more fundamental circuit element than the *D*-flip-flop in that edge-triggered behavior is attained only by implementing two latches and generating a pair of two-phase nonoverlapping clock pulses internally, in response to the active clock transition at the edge-triggered input.

For instance, Figure 9.38 shows a typical *D*-flip-flop in which inverters I_1 and I_2 generate the internal two-phase clock signals for level-sensitive latches L_1 and L_2. In specialized applications, it may be advantageous to design a custom module within which multiphase clocks are distributed directly, without incurring greatly increased skew problems. For example, an error-correcting codec ASIC prototype for 45 Mb/s digital communications includes a 2613 stage tapped delay line comprised of seven-gate single-phase *D*-flip-flop modules. The use of single-phase clock flip-flops in this context is relatively expensive, but appropriate for fast validation of the system design. For cost- and power-reduced production in volume, a more detailed latch-based design using directly distributed two-phase nonoverlapping clocks may be worthwhile. In general, while it is most common to conduct system design based on the single-phase clocking model, two-phase or multiphase clocking may be advantageous within specialized substructures.

Although edge triggered, a minimum clock pulse width is still typically required to deliver enough switching energy on the clock line. For correct operation (specifically, to avoid uncertain outputs due to metastability) of an edge-triggered flip-flop, data must be stable at the input(s) for a minimum setup time before the clock edge, and the data must remain stable at the input for the hold time, after the clock edge. The time until the *D*-flip-flop output is valid after the clock edge occurs is the clock-to-*Q* delay. For hazard-free transfer of data from one stage to another with *D*-flip-flops, without assuming a minimum logic delay constraint between stages, the clock-to-*Q* delay must exceed the hold time. Typical values for a range of *D*-flip-flop types in a 1.5 μm CMOS process are $t_{\text{setup}} = 0.8$–1.5 ns, $t_{\text{hold}} = 0.2$–0.4 ns, and $t_{\text{clk}-Q} = 1.3$–3.5 ns for *Q* output fanouts of 1–16, respectively. With the extra input logic delays in a JK flip-flop, many JK flip-flop cell implementations exhibit $t_{\text{hold}} = 0.0$ ns. By comparing magnitudes of typical setup and hold time requirements, it is apparent that skew is more likely to cause a setup time violation on critical delay logic paths than it is to result in a hold time violation.

9.2.2.11 Role of Clocks in Dynamic Logic

Clock signals are also used to implement a variety of logic gate functions in a dynamic circuit style, i.e., based on short-term charge storage, not static logic. This typically involves precharging on one phase and logic evaluation steered by the inputs on the second phase. The ''Domino'' logic approach combines a dynamic NMOS gate with a static CMOS buffer [8]. In ''NORA'' (no-race), logic dynamic logic blocks are combined with clocked CMOS latch stages. A variety of other dynamic logic circuits, using up to four clock signals to structure the precharge and to evaluate timing, are covered by [12,31]. In all of these gate

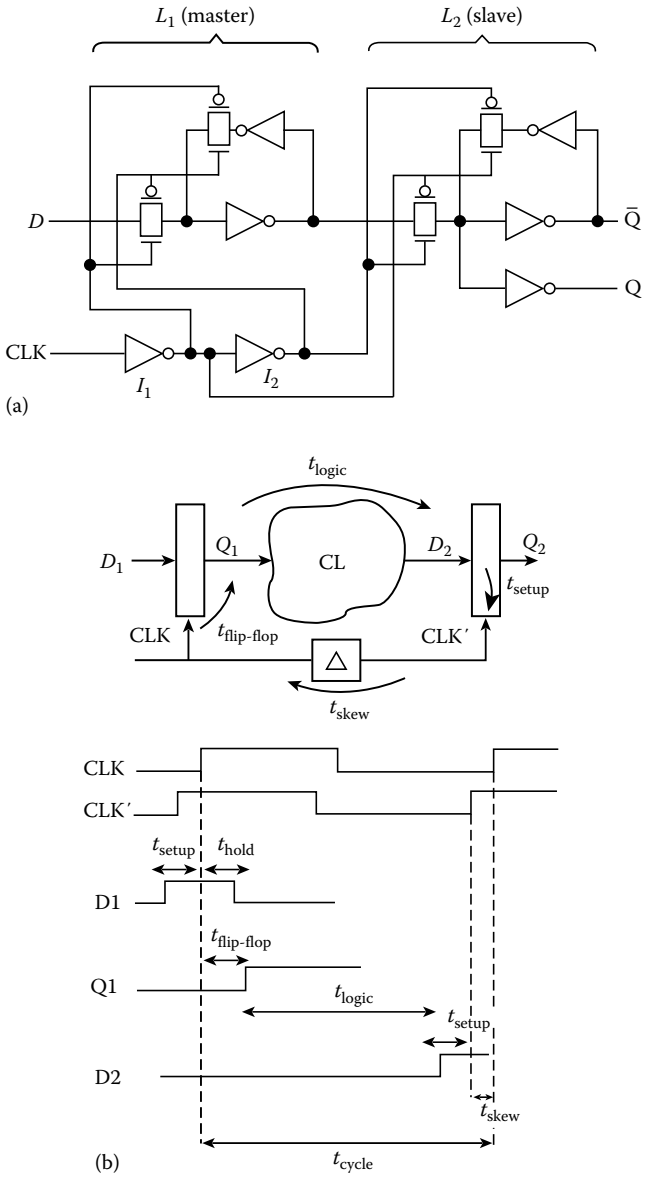


FIGURE 9.38 (a) Two-phase latch structure of a typical CMOS positive edge-triggered *D*-flip-flop; and (b) setup, hold, and delay times for a *D*-flip showing how skew is equivalent to a shorter clock period and threatens setup time margin. (From Bakoglu, H.B., *Circuits, Interconnections and Packaging for VLSI*, Addison-Wesley, Reading, MA, 1990, 345.)

level circuit implementations, the clocking-related issues are ultimately manifestations of the basic principles already seen for two-phase clocking; i.e., of never simultaneously enabling a direct path from input (or precharge source) to output. These logic styles were developed to reduce transistor counts. However, modern designers will most often be faced with a greater challenge in managing system-level problems of skew in a single clock phase distributed to static register than the challenge of reducing transistor count.

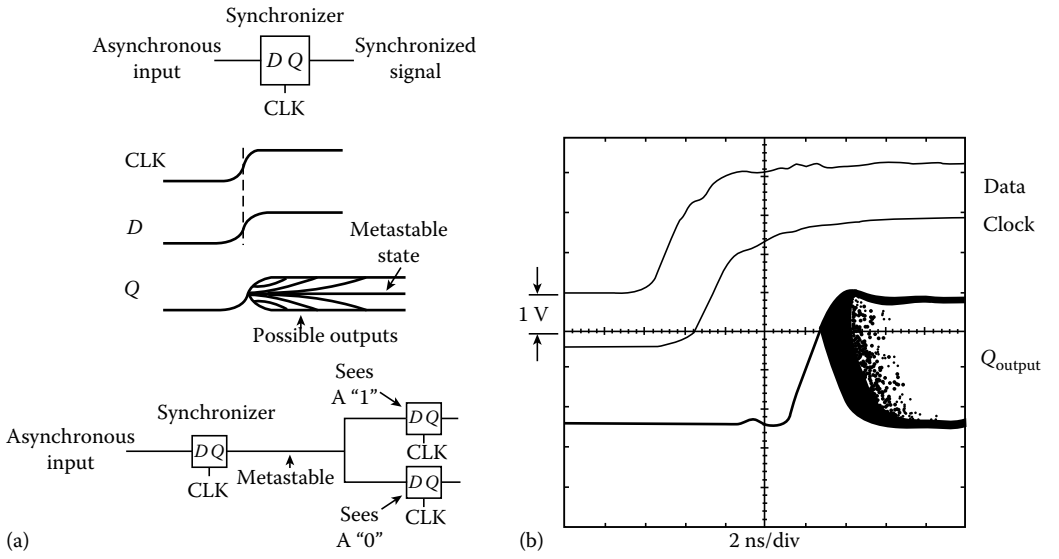


FIGURE 9.39 (a) Metastability in a synchronizer circuit. (From Bakoglu, H.B., *Circuits, Interconnections and Packaging for VLSI*, Addison-Wesley, Reading, MA, 1990, 357.) (b) Experimental illustration of metastability in 74F/74 (TTL) *D*-flip-flop showing output rise before indeterminate final response (10 s point accumulation). (From Johnson, M.W. and Graham, M., *High-Speed Digital Design: A Handbook of Black Magic*, Prentice Hall, Englewood Cliffs, NJ, 1993, 130.)

9.2.2.12 Synchronizers and Metastability

Many systems need to sample external inputs which may be timed independently of the synchronous system clock, such as switch-based control inputs, keyboard states, or external process states in a real-time controller. The external state needs to be synchronized with the system time base for processing. Metastability leading to synchronizer failure is a fundamental possibility that can never be entirely eliminated.

Figure 9.39 is a basic synchronizer circuit. Our concern is that it is possible for the synchronizer output to take an arbitrarily long time to settle to one or the other valid logic states if the input signal voltage is sampled in the intermediate voltage range, i.e., $V_{iL} < V_{in}(t) < V_{iH}$. In this range, it is possible to find the input voltage at a value that leaves the cross-coupled latches internal to a flip-flop in an intermediate state, with insufficient positive feedback to snap the output to either high or low valid states. System noise or quantum fluctuation will ultimately perturb such a precarious balance point and the output runs to one direction or the other, but it can take an arbitrarily long time for the synchronizer to reach a valid output state. As shown in Figure 9.39b, some flip-flop outputs may also tend to rise at least halfway toward the positive output level before deciding the input was really a zero. This glitch may trigger edge-sensitive circuits following the synchronizer.

Fortunately, the probability of an indeterminate latch output falls exponentially with the time T after sampling the possibility indeterminate input

$$P(t > T) = f_{clk} f_{in} \Delta e^{-T/\tau_{sw}} \tag{9.5}$$

where

f_{clk} is the synchronous sampling frequency
f_{in} is the frequency of external transitions to be synchronized
Δ is the time taken for the input voltage in transition to cross from V_{iL} to V_{iH} (or vice versa)
τ_{sw} is the time constant characterizing the bandwidth of the latch device

Having recognized the strict possibility of a metastable logic state resulting from synchronizer input, the designer can address the issue in several practical ways:

1. Use a high gain fast comparator to minimize Δ by minimizing the voltage range V_{iL} to V_{iH}.
2. Ensure or specify fast transition in external sensors or other devices to be sampled, if design control extends to them.
3. If there is no real-time penalty from an additional clock period of input response delay, the synchronizing latch should be followed by one or two more identical synchronizer latch stages, thereby increasing T in Equation 9.5 to reduce the chance of a metastable state being presented to internal circuitry to an acceptably low probability.

The effect of input metastability on the system also should be analyzed for its impact. If it is extremely crucial to avoid a metastability hazard, then phase-locking the external system to the system clock may be considered, or if the system and external timebases are free running but well characterized (e.g., in terms of a static frequency offset or known phase modulation), then the anticipated times of synchronization hazard may be calculated and avoided or otherwise resolved.

As a practical matter, the way in which design software handles metastability should be considered. Potentially metastable conditions should be flagged as a warning to the user, but not necessarily treated as a violation prohibited by the design environment. Some applications, particularly in VLSI for telecommunications, need design support for plesiochronous (near-synchronous), phase-modulated, or jittered signal environments. This implies test vector support to represent clocks interacting through logic at slightly different long-term or instantaneous free-running frequencies with design and simulation rules that permit the metastable conditions inherent as such clocks walk relative to one another. Circuit simulations must be allowed to continue with a random value resulting from the simulated ''synchronizer failure'' to be useful in such applications.

9.2.2.13 Controlled Introduction of Skew

Skew is not necessarily all bad. In fact, from the viewpoint of the system power supply, and power and grounded-related noise current surges, it is undesirable to have all logic transitions occurring exactly simultaneously. In a CMOS IC with 20 K register stages at 0.1 pF load each and a 1 ns clock rise time, 10 A of peak current can be drawn by the clock net. This can present a serious $L\,dI/dt$ problem through power and ground pins and can even lead to electromigration problems for the metallic clock lines. Chip clocking strategies should take this into account early in the design by seeking ways to deliberately stagger or slightly disperse the timing of some modules with respect to others. Also, the system clock waveform may not necessarily need the fastest possible rise time. Consistent with avoiding slow-clock problems, and controlling threshold-related skew in buffers, the clock edge should not be made faster than this as an end in itself. Excessively fast clock edges clock edges only aggravate power and ground noise problems as well as ringing and potentially causing electromagnetic radiation problems in the chip-to-chip interconnect. These principles motivate the widely used 10 K ECL logic family, which is based on the much-faster 100 K series with explicit measures to slow down the rise and fall times of the basic 100 K logic gates.

When considering random skew, it may or may not be beneficial to pursue skew reduction below a certain level in the design. In the case of a microprocessor or a design in which the fastest possible IC speed is always useful skew reduction does mean a performance improvement. In some other applications, however, the clock speed is set by the application. For instance, a VLSI circuit for a telecommunications MUX may be required to operate at a standard line rate of 45 MHz. In this case there may be no premium for a design that can perform the same functions at a higher speed. A working design with skew of 5–7 ns (out of a 22 ns clock period) may then be more desirable than a functionally equivalent design with 0.5 ns skew because dI/dT effects are eased by distributing the total switching current over time in the former. This principle may be important in practice as automated clock synthesis tools become more widely used and effective at achieving low skew, possibly creating unnecessary system-level noise and

EMI emission problems. Future clock-related CAD tools should possibly aim to disperse clock timing at various loads while satisfying a target worst-case skew, rather than absolutely minimizing skew.

Strictly speaking, skew can also be beneficial when allowed to build up in a controlled way in certain regular logic structures. For instance, by propagating clock in the opposite direction to data in a shift register, one enhances the effective setup time of the data transfer from register to register. In general, however, it is not feasible or advisable to try to design every clock path with a desired (nonsimultaneous) clocking time at each register, taking into account the individual logic paths of signals leading each clocked latch input. Especially when designing larger systems mediated by CAD tools for placement, routing, and delay estimation, the most practical and low-risk approach is to consider any deviations from a common nominal clock time as undesired skew. Indeed, for any one latch, timing margin may be enhanced by the actual skew that arises, but with thousands of logic paths, it is impossible to analyze the relative data and clock timing for each latch. Only one instance in which the skew works against the assumed timing margin is enough to fail a design. Therefore, the ''customized skew'' approach is recommended only for small and very high speed specialized circuit design.

9.2.2.14 Clock Signal Manipulation

As a matter of design discipline, some commercial ASIC and cell-based layout systems may prohibit a designer from directly gating or manipulating a clock signal. Any needed clock qualification is done through defined enable or reset inputs on register structures. As in software development, in which structured design disciplines have been developed, gating the clock may be riskier than its apparent efficiency warrants. In addition, clock gating within random logic designs can interfere with test pattern generation. The risk also exists of creating clock glitches or even logical lockups when clocked logic functions decode conditions that gap its own clock. On the other hand, in high performance and in large system-level designs, clock gating for power down and ''clock tuning'' may be unavoidable.

Wagner [30] discusses clock pulse-width manipulation, defining four canonical subcircuits that can be used to ''chop,'' ''shrink,'' or ''stretch'' the clock waveform for either delay tuning or duty cycle maintenance. The effect of these circuits on the positive pulse portion of a clock waveform is shown in Figure 9.40, where AND gates have delay d_a, OR gates have delay d_0, inverters have delay d_i and the delay elements have delay D. Aside from a single gate delay, the chopper and stretchers leave the rising edge unaltered and tune the trailing edge. These can be used to maintain a balanced clock duty cycle or to tune the nonoverlap period in two-phase clocking. The shrinker delays the rising edge of the clock as might be helpful to specifically delay clocking a latch or a flip-flop that is known to follow a particularly long logic path delay. This is not generally a preferred design approach, especially when manufacturing repeatability and temperature dependence of delay elements are considered.

By ''clock gating,'' we mean selectively removing or masking active phases or edges from the clock signal at one or more latches. One valid reason to gate the clock in CMOS is to reduce power consumption. Many circuit designs possess large modules or subsystems which it makes sense to stop cycling in certain application states. Gating the clock off is therefore the simplest form of power management, because CMOS has negligible power dissipation in a static state. However, even for this simple use of clock gating, the main issue is avoiding glitches when gating the clock.

Before gating any clock, the designer should see if gating can be avoided with an alternate design style. For example, if it is desired to hold the data on one register for a number of cycles while other registers on the same clock proceed, a preferred approach is to use a 2:1 MUX at the register input. Rather than gate the clock, the MUX is steered to select the register's own output for those cycles in which gating would have occurred. Ultimately, if it is appropriate to gate out one or more clock pulses, a recommended way of doing so in rising edge active logic is to OR out the undesired clock edges, decoding the clock gapping conditions on the same clock polarity as the one being qualified (see Figure 9.41). A natural tendency seems to be to AND out the gapped clock edge and/or to decode the gapping condition on CLK, but these approaches are more apt to generate clock line glitch than the OR-based approach. In the AND approach

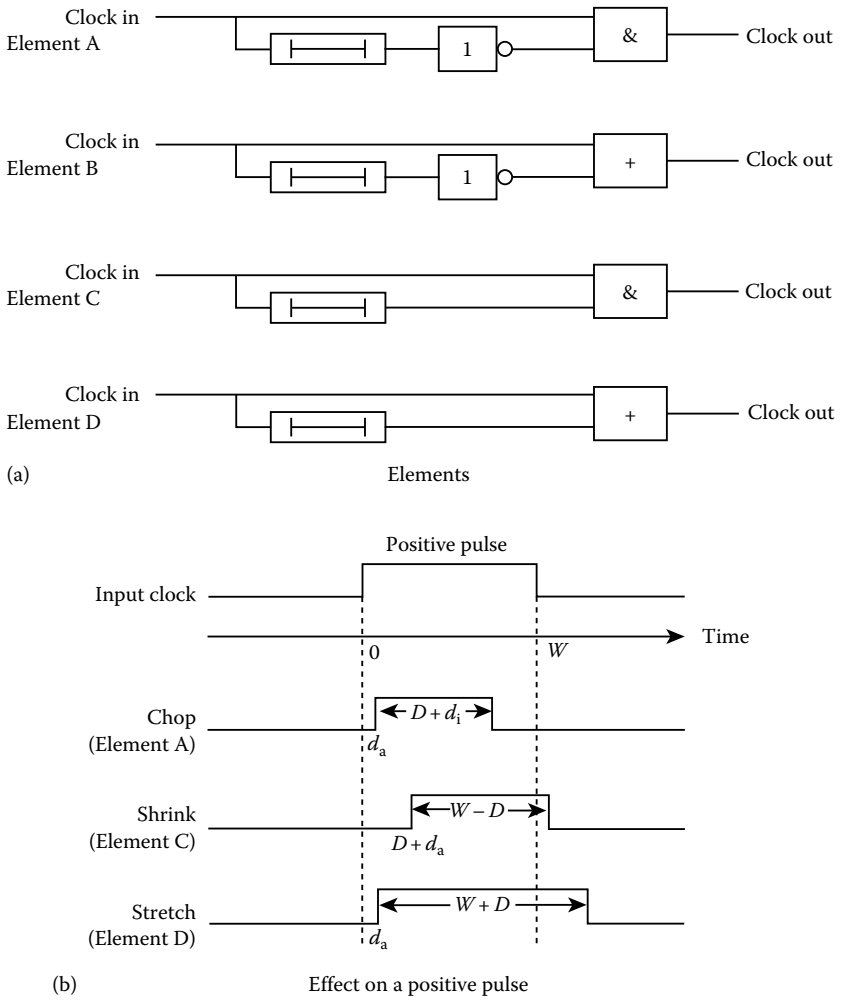


FIGURE 9.40 Standard circuits for chopping, stretching, and shrinking a clock waveform to adjust duty cycle or timing margins. (Adapted from Wagner, K.D., *A Survey of Clock Distribution Techniques in High Speed Computer Systems*, Report CRC 86-20, Stanford University Center for Reliable Computing, Stanford, CA, December, 1986, 15.)

the gating line returns high at the same time as the falling edge after the last-gapped active edge. In the case of minimum delay through the gapping logic the risk is that both AND inputs are momentarily above threshold.

9.2.2.15 Minimizing Delay Relative to an External Clock

In a large system, skew can build up between clock and data at the system level, even if the clock is skew-free everywhere within the ICs because data paths through ICs can build up delay relative to the system clock. For instance, if an ECL or TTL system clock line is distributed to a large CMOS IC, then the system clock must be level shifted and a large clock buffer may be required in each IC to drive its internal clock net. The delay through the on-chip clock interface and buffer can mean that even if the chip timing is internally skew-free, the on-chip clock is significantly delayed relative to the external system timing. Short of using the phase-lock methods described later, a simple technique to minimize this form of system-level skew is either to retime chip outputs with a separate copy of the system clock that has not

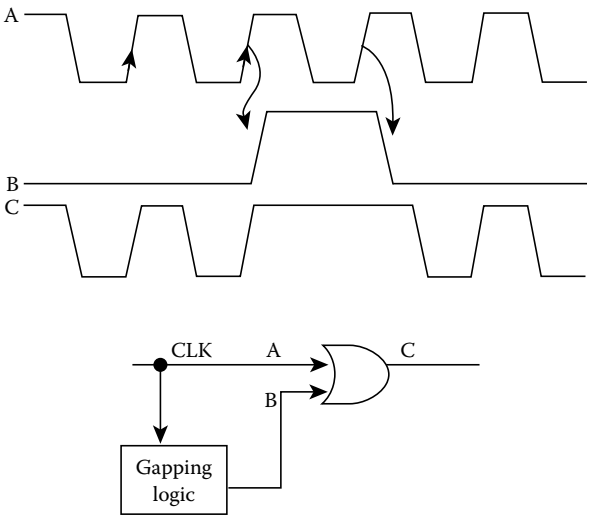


FIGURE 9.41 OR-ing out a clock edge when clock gating is essential.

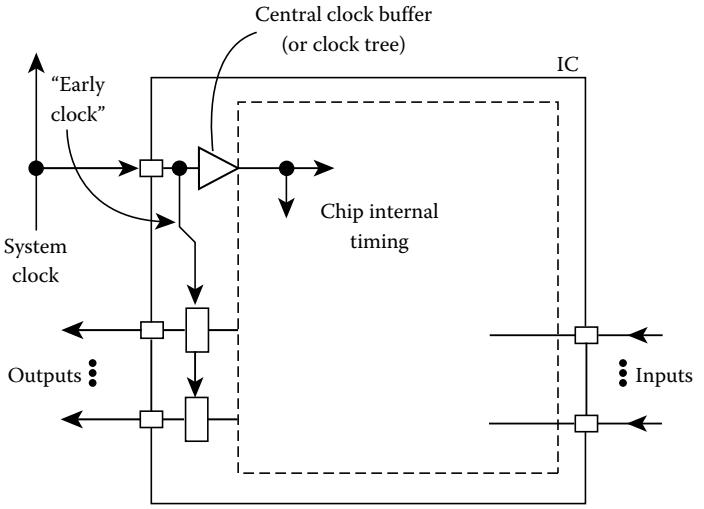


FIGURE 9.42 The "early clock" technique for reducing chip delay relative to external system timing.

gone through the internal clock buffer, or, if electrically compatible, to use the external system clock to directly retime the output signals from each IC (Figure 9.42). This is called the "early clock" concept. Note that this assumes an adequate timing margin exists in the final stage of internal logic to permit the relatively early sampling of logic states.

9.2.3 Clock Distribution Schemes

9.2.3.1 Single-Driver Configurations

Often a single on-chip clock buffer is the simplest and best approach to clock distribution. A single adequately sized clock buffer is typically located at the perimeter to drive the entire clock net of the chip,

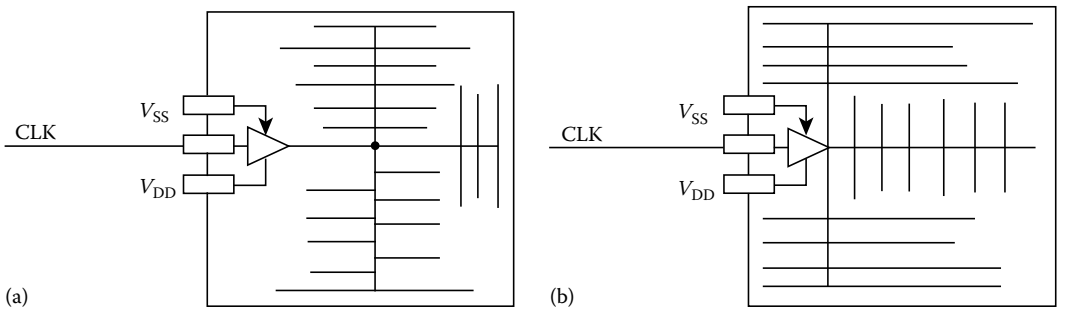


FIGURE 9.43 Single clock buffer placed in the I/O perimeter with dedicated power, ground pins (a) branching from a medial point on die (current density on line to medial point may be high) and (b) branching immediately (skew may be high).

as shown in Figure 9.43. This approach can also perform well in large systems if the clock is distributed in a low R_0C_0 routing layer, such as third-layer metal. The main advantage regarding skew is that no matter how the single-clock driver delays or otherwise responds to the external clock input, the output waveform is electrically common to all loads. No intervening buffers and no separate passive net segments develop skew. Moreover, if the global clock net comprises an electrically isochronic region ($NC_{gate} \gg l^2\,RC_{wire}$) in which clock loads are reasonably uniformly distributed, the clock net voltage rises virtually simultaneously at all points on the charging clock net, resulting in extremely low skew. Often this leads to lower skew than in a buffered clock fanout tree. There is also only one global clock wiring net, simplifying documentation.

On the other hand, skew can be larger and more routing dependent with a single buffer (than with some following schemes) when clock routing lengths vary significantly and wiring capacitance and resistance are significant. In such cases an isochronic net is not an accurate model, and performance depends on the way clocked loads are distributed on arms branching from the medial clock node. It is possible for neighboring flip-flops to be connected to the central driver via quite different path lengths, making prelayout simulation relatively uncertain and requiring considerable postlayout tuning. Another caution is that even if no actual skew is present because a single waveform is common to all loads, the rise time of the clock waveform may show considerable loading effects, so that threshold-dependent skew arises in the clocked loads. Finally, the potential for conducted switching noise problems is high with the single-buffer configuration because the entire clock net capacitance is charged in typically under 1 ns. Power supply decoupling and ground bound problems and even current density (electromigration limits) considerations may need to be given special attention if using this configuration. It is usually recommended that the single clock buffer be physically adjacent to the clock input pin, and the clock pin should be flanked by dedicated power and ground pins that feed only the clock driver. This principle, which applies in general to clock input buffers in all clocking schemes, keeps the clock switching noise out of the core power and ground bus lines. Also, the delay through a large clock buffer may be considerable, so a lightly loaded ''early clock'' can be picked off from the buffer input and used to retime chip output registers, or a PLL may advance the phase to the internal clock buffer to align internal and external timing regardless of the buffer delay.

An interesting central clock buffer design, which also has attributes of the clock trunk scheme which follows is reported in [9]. Here, the area for central clock driver fabrication is a strip across the center along one axis of the die. External clock is fed from one pin on the same axis as the buffer, and internal clock lines radiate systematically away from the central linear distributed buffer. Data flow is also highly structured and arranged to progress away from the central driver strip, further minimizing clock-to-data skew.

9.2.3.2 Four-Quadrant Clocking Approaches

In the quadrant-oriented approach, we may use up to four clock pads and four smaller clock drivers placed at the centers of the die edge, preferably also with dedicated power and ground from flanking pins. There may be one, two, or four external clock pins. Figure 9.44a shows a quadrant scheme tested in [24]. In Figure 9.44b, a single-pin quadrant-oriented scheme in 1 μm, two-layer metal CMOS achieved 0.6 ns skew among 400 registers. A four-pin quadrant approach was successfully used in [29] to develop a 90 MHz CMOS CPU. If more than one pin is used for clocking, pin count goes up but absolute delay through the clock buffers can be reduced. The maximum wiring RC delay on the internal clock net and the peak current and $L\,dI/dt$ effects through any one pin and bonding wire inductance may all be reduced in this case. Total loads on each of the four internal clock nets should be well balanced and/or the drivers sized for the actual number of loads in each quadrant. In many cases, reduction of each clocked area to one fourth of the die area can result in isochronic regions for which no further design attention other than total load balancing is required for clock routing within each region. The quadrant approach can also reduce the clock routing problem when the clock shares only two metallization layers available for all signal routing. Two other considerations apply to the quadrant schemes: (1) external tracking to multiple clock pins should be laid out with delay balancing in mind; (2) skew should be particularly considered on data paths which must cross from quadrant to quadrant, bridging timing across clock subnetworks.

9.2.3.3 Symmetric and Generalized Clock Buffer Trees

A symmetric or regular clock buffer tree (Figure 9.45a) has equal fanouts from buffers at the same level, equivalent (or isochronic) passive interconnect paths at each stage, identical buffer types at each level, and equal groups of loads at each leaf of the tree. This ideal can be approximated if loads are regularly distributed and layout and routing can be controlled to ensure equal interconnect delays from each buffer to its subtending buffers. The remaining skew will primarily be due to threshold variation in buffers and terminal loads.

A more general view of the clock buffer tree that arises with irregular layouts and routing is that the buffer tree has unequal interconnect delays and differing numbers of loads at each buffer. Figure 9.45b illustrates the electrical model of such a clock tree. (R and C values are all different.) The basic approach to skew control in such a generalized buffer tree is to size each buffer specifically for the individual loads it drives and the delay of the interconnect path to its input from the preceding level buffer. In practice, this means that generalized clock buffer trees may be the most handcrafted (and/or custom-tuned) of all designs, especially in a large system-level clock tree that extends down from a master clock source through multiple shelves, backplanes, connectors, circuit packs, and individual ICs. The system-level hierarchical clock tree design for the VAX 8800 is a good example described by Samaras [25]. Here, a two-phase clock was distributed to 20 large circuit packs over a 21 in. backplane with a global skew of 7.5 ns, for operation at about 50 MHz (37% skew).

Some basic methods and principles are, however, identifiable for designing generalized buffer trees so that it is not all just careful tuning and handcrafting:

1. Inverting buffers at each level will preserve clock duty cycle better than a tree of noninverting buffers.
2. Total delay (root to leaves) of the clock net is theoretically minimized when driving primarily capacitive loads, by a fanout ratio $e = 2.718\ldots$ at each level of the tree [23]. In practice, this implies a fanout ratio of about $n = 3$, with appropriately sized buffers, if delay is to be minimized. However, $n = 3$ can lead to relatively deep trees of many buffers in which skew can build up from threshold and time-constant variations as in Equation 9.4.
3. If identical buffers are used within each level, then the design aim is to make sure that the load of further buffers and/or interconnect RC load is delay-equivalent for each buffer. Dummy loads may be required to balance out portions of the clock tree whose total fanout is not needed. At the end-nodes of the tree equal numbers of standard loads should be grouped together on each leaf node, within a locally isochronic region.

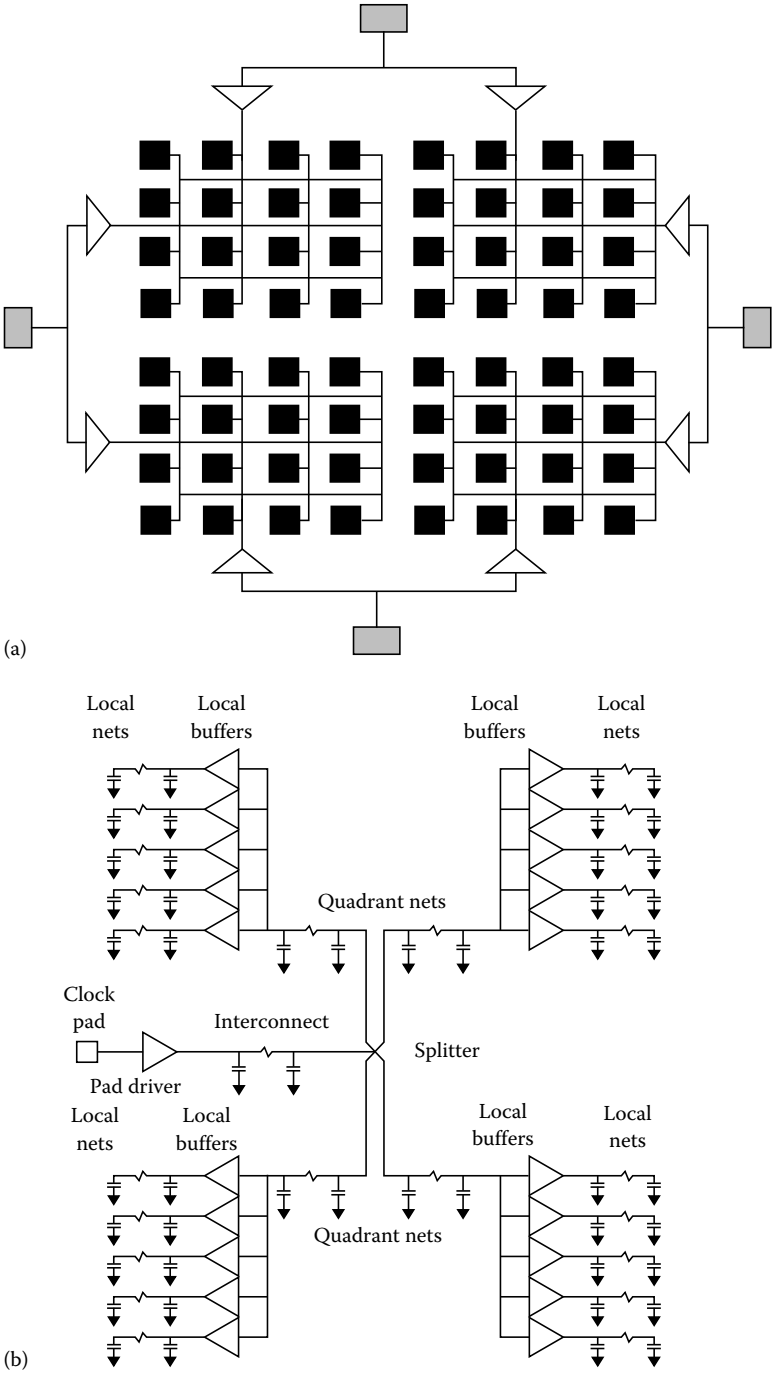


FIGURE 9.44 Quadrant-oriented clock distribution schemes: (a) 4 pins, 2 parallel buffers per quadrant. (From Nigam, N. and Keezer, D.C., A comparative study of clock distribution approaches for WSI, *Proc. IEEE 1993 Int. Conf. WSI*, 243–251, 1993.) (b) Single-pin quadrant scheme with 2 buffer levels. (From Boon, S., Butler, S., Byrne, R., Setering, B., Casalanda, M., and Scherf, A., High performance clock distribution for CMOS ASICS, *Proc. IEEE 1989 Custom Integrated Circuits Conf.*, 1989, 15.4.1–15.4.5.)

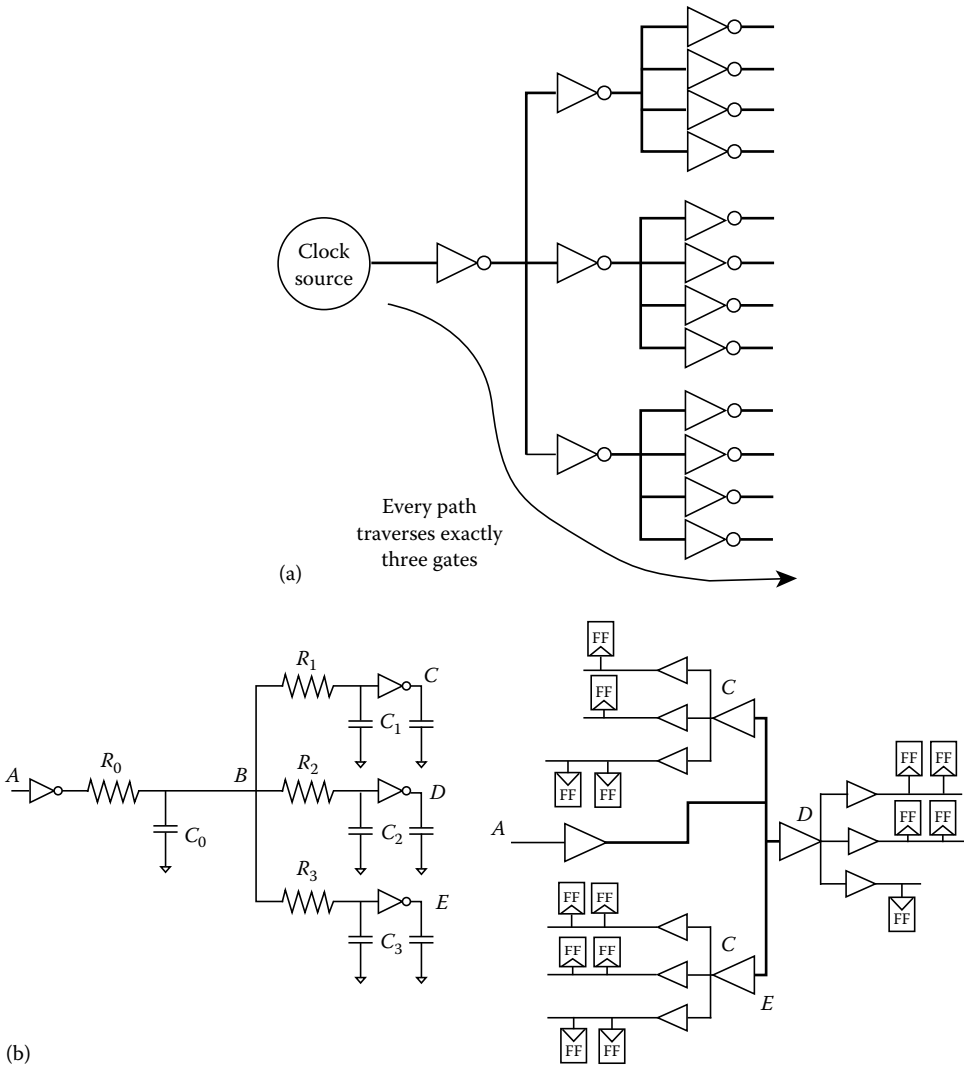


FIGURE 9.45 (a) Idealized symmetric buffer clock tree. (From Johnson, M.W. and Graham, M., *High-Speed Digital Design: A Handbook of Black Magic*, Prentice Hall, Englewood Cliffs, NJ, 1993, 348.) (b) Generalized clock buffer tree where interconnects and loads are not identical.

4. If the tree is deep, the skew among members of one local clock group at the bottom of the tree may be considerably smaller than the skew than the skew between groups of clocked loads at different end-nodes of the tree. If the inter- and intragroup skews are separately characterized, however, logic path delays can be designed to take advantage of the low intragroup skew and to allow greater timing margin on intergroup paths [25].
5. The choice of clock buffer type for use in any clock tree should take into account the effects of power supply sensitivity. For example, the "bootstrapped clock buffer" of Figure 9.46a can provide a very sharp rise time, although with relatively high delay through the buffer. Sharp rise times minimize skew due to switching threshold variations in the following buffers or clocked loads. The output switching time of the bootstrapped clock buffer is, however, relatively sensitive to supply

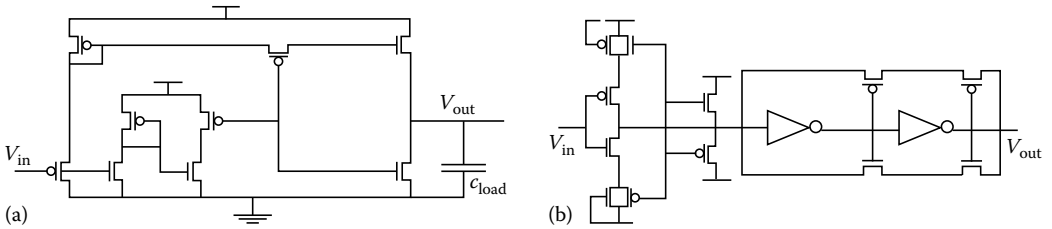


FIGURE 9.46 (a) Bootstrapped clock buffer for use in first level of a clock tree; and (b) phase-correcting clock buffer for use deeper in clock tree. (From Fried, J., *Proceeding of the IFIP Workshop on Wafer Scale Integration*, G. Saucier and J. Trilhe, Eds., North-Holland, Amsterdam, 1986, 127–141.)

voltage. On the other hand, the ''phase-correcting buffer'' of Figure 9.46b is very tolerant to supply variations, but is not as fast in its output rise time. This leads to a mixed buffer strategy in which the bootstrapped buffer is used in the relatively small population of buffers in the first few stages of a clock tree. Here, special attention can be paid to ensuring well-equalized power voltages. The phase-correcting buffer is more appropriate in later stages, nearer the more numerous individual loads, among which on-chip or wafer supply levels may exhibit more IR voltage drop variation.

Algorithms for the generalized clock tree design problem are also emerging. The algorithm in [28] can generate a buffer design based on custom transistor sizing to drive each heterogeneous load, all with the same nominal delay. A still more general optimization approach is under development by Cirit [5]. The sizes of all buffers in the tree are jointly optimized with respect to unequal interconnect RC totals and unequal loads that each drives, as in Figure 9.45b. Thus, the minimum ''total'' tree delay is found for which all paths from root to leaf of the tree also have ''equal'' delay. The procedure is intended to be iterated along with placement and routing alternatives until a set of feasible buffer sizes and acceptable total delay is found.

9.2.3.4 Clock Trunk Schemes

The clock trunk concept is gaining popularity and is now supported within several CAD systems for CMOS processes with two or more metallization layers. Three variants of clock trunk structures are shown in Figure 9.47. An input clock signal is buffered (its input pad is at the center of one side of the chip edge) and is routed either to the midpoint, or one or both ends of the internal clock ''trunk.'' The trunk itself is a metal line specially widened for low resistance, thereby making delay and the $R_0[C_{load} + C_0]$ rise time particularly small on the trunk portion of the overall clock net. As long as the lumped capacitive loads (C_{load}) dominate the trunk's C_0, the time constant of the trunk drops as it is widened. C_0 can be kept particularly low by forming the trunk in third-layer metal. The idea is to size and drive the clock trunk so that an isochronic region is created down the middle of the die, reducing all remaining clock path distances to not more than half the die diameter and setting up a situation in which the final distribution of clock from the isochronic trunk to the loads is in line with one routing axis. This means that branch routing from the trunk to loads can be exclusively contained within one metal layer and travel in a shortest direct line from trunk to load. This is highly desirable for processes in which all horizontal and vertical routing are dedicated to one metal layer or the other. Overall layout and routing is simplified and clock paths are predictable and uniform in routing and wiring type. If, however, the total fanout is very small, the wide metal trunk may add more to trunk capacitance than it decreases R_0 in the $R_0[C_{load} + C_0]$ product for the trunk. This is undesirable because for an isochronic trunk, we want C on the trunk to be dominated by its loads, not by the distributed capacitance of the trunk. In practice, therefore, the trunk scheme is typically recommended for fanouts of over 50. Below this, a single buffer scheme is recommended.

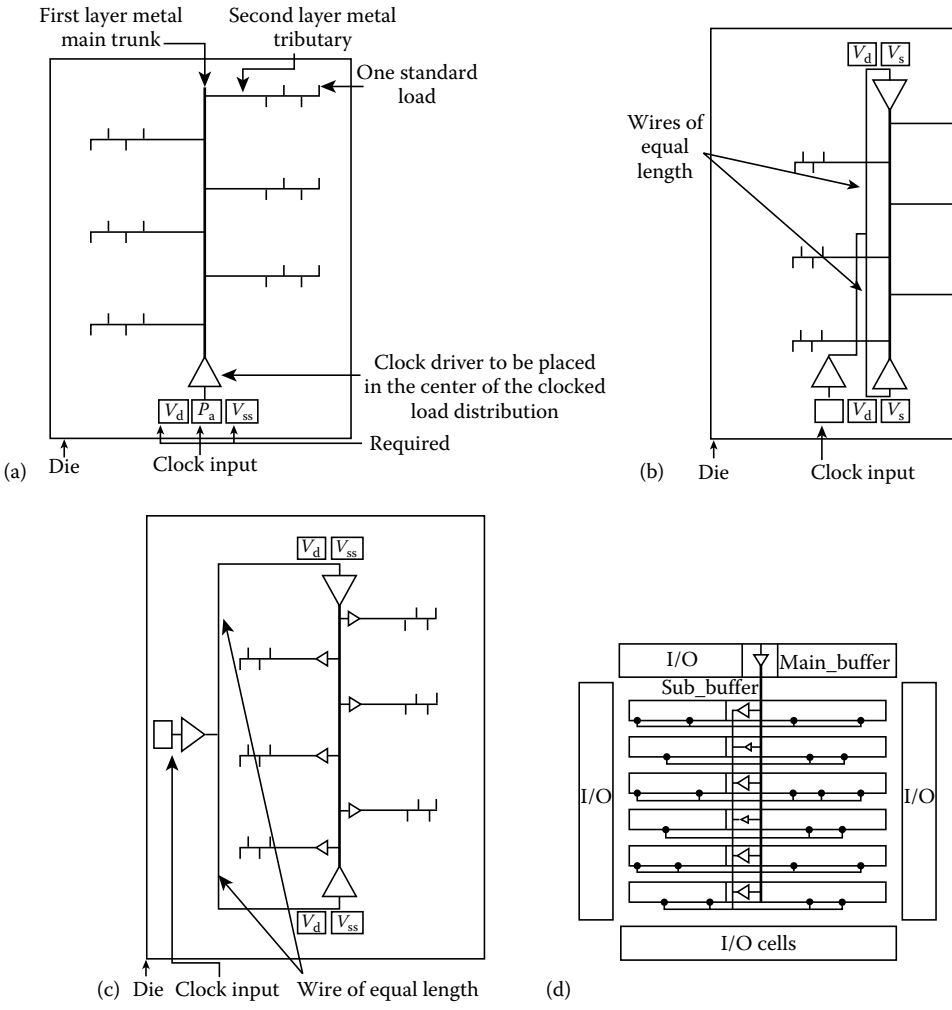


FIGURE 9.47 The clock trunk concept: (a) single-ended unbuffered clock trunk, (b) double-ended unbuffered, (c) buffered clock trunk, (From LSI Logic Corp., *Clock Scheme for One Micron Technologies*, Rev. 1.1, LSI Logic Application Note, Aug. 1992.) and (d) clock trunk with shorted branch buffer outputs. (From Saigo, T., Watanabe, S., Ichikawa, Y., Takayama, S., Umetsu, T., Mima, K., Yamamoto, T., Santos, J., and Buurma, J., *Proc. IEEE 1990 Custom Integrated Circuits Conf.*, 1990, 16.4.1–16.4.4.)

By designing with the following principles for single-ended, double-ended, and buffered clock trunks, clock nets of up to 2000 fanouts can achieve $<$1.5 ns skew in 1.0 or 0.7 μm CMOS gate array technology [21]. When the clock fanout is between 50 and 500 unit loads, a single-ended trunk scheme, as shown in Figure 9.47a provides a good trade-off among skew, area, and delay. A single clock driver input buffer is used to drive the trunk line, which is typically realized by six first-layer metal lines in parallel with metal filled in. Clock trunk sizing for a given fanout must set a minimum width to take current density limits into account, given that all of the clock net current flows through the trunk if the tributaries are not buffered. Tributaries of nominally constant fanout branch out in second-layer metal. To the extent possible, macrocells and hard-coded megacells should be laid out with the clock trunk in mind, ideally permitting the clock trunk to be located in the middle of the logic to which it fans out.

The tributary branches may or may not be buffered with their own clock drivers, depending on the total number of loads to be driven. Buffering primarily has the effect of reducing overall delay, rise time, and total current density in the trunk, by allowing a smaller trunk driver. Form a purely skew-oriented view, however, it is better not to have the secondary drivers as long as the trunk is isochronic and the size and delay of the main driver is acceptable. When using local buffering, it is important that the branch loads on all tributaries be balanced, more so than when using an unbuffered trunk. Local buffering is, therefore, primarily a way of distributing the total buffering load so that no one buffer needs to be extremely large.

For layout software simplicity, the main trunk may be constrained to use vertical (or horizontal) routing channels only. The main trunk (or each of possibly several main trunks) should be placed as close as possible to the centroid of area of all the loads which it drives. Layout or floorplanning software for commercial ASIC design can typically assist the designer in clock trunk placement by visualizing the spatial distribution of clock loads.

When a design has between 500 and 2000 clock fanouts, a double-ended clock trunk, as in Figure 9.47b or c is recommended. The double-ended clock drivers are internal buffers which use the I/O slots of two pins that will not be used externally thereafter. Both single-ended and double-ended clock trunk schemes use only one external clock input pin, with adjacent pins providing an AC ground and switching noise isolation by powering the drivers with dedicated V_{ss} and V_{DD} pins for the clock buffers. In general, clock input pins should always be surrounded by nondriven pins to minimize the possibility of cross talk coupling into the clock waveform. Clock pins also should be chosen so that minimal internal routing is required between the predriver associated with the input pad and the clock trunk drivers. As a clock trunk design is laid out, the spike current draw from V_{ss} due to simultaneous switching of large fanouts on the clock net should be assessed and considered in determining how many V_{ss} pins are needed in a particular design.

In the double-ended trunk scheme, some care must be taken to ensure that an equal-length path can be routed from the midpoint of the trunk, where the clock input branches to the trunk drivers at both ends of the clock trunk, and that a direct routing from the side of the die to the branch point at the center of the die is feasible. Particularly when preconfigured macrocell function have been placed, the metallization layer needed to bring the clock predriver into the middle of the die may be blocked. This leads to the recommendation that the clock input pin be placed on the side of the die that is in line with the clock trunk (see Figure. 9.47c). The line to the branch point and the two branch lines can then use the same metallization layer as the clock trunk and can be automatically provided for as part of the routing channel width reserved by the clock trunk layout software.

The two double-ended arrangements will have a basic skew given by the (nontrunk) R_0C_0 delay across one half of the chip's dimension (typically under 300 ps), plus skew due to any imbalance in buffer loads and thresholds. An advantage of the buffered clock trunk is that the capacitive load of the clock tree is distributed somewhat in both time and position across multiple buffer stages, reducing the current spikes occurring during a clock edge and their impact on ground bounce and injected power supply noise. On the other hand, a total of three or four buffer stages associated with this structure (for low skew in large applications) may cause high delay between the clock edge used to latch incoming and outgoing data on chip and the external system clock. Clock net fanouts achievable using the clock trunk scheme in commercial gate arrays are summarized in Table 9.10.

In even larger dies with famous of over 3 K flip-flops, multiple symmetrically driven double-ended clock trunks can be established to control the maximum distance of any point from a clock trunk. For example, with two trunks placed one fourth of the die width in from the sides, no load is over one fourth the die diameter from a trunk, and the loading of branch buffers is half that required with one trunk. Branch lines from different trunks should not be connected together where they meet in the middle of the die. These points are far enough away in terms of delay from their common driving points that joining them could cause power-wasting buffer output fights. In a further variant on the buffered clock trunk, buffers have their outputs ganged (i.e., shorted) by an additional vertical metal line parallel to the trunk,

TABLE 9.10 Fanouts of Clock Trunk Schemes in 1 μm CMOS Gate

Clock Frequency (MHz)	Single-Ended Trunk	Double-Ended Trunk	Double-Ended Trunk with Local Buffering
50	500	2000	>2000
60	450	1500	>1500
70	400	1200	>1200

Source: LSI Logic Corp., *Clock Scheme for One Micron Technologies*, Rev. 1.1, LSI Logic Application Note, Aug. 1992.

Note: Arrays (maximum number of clock loads driven with <1.5 ns skew).

close to the buffer outputs. The effect of shorting the branch buffers is to equalize the propagation delay through the trunk and distribute the capacitance per buffer more uniformly. This has been found to reduce skew considerably if the branch buffers were not equally loaded and also reduced skew (although less so) in the balanced buffer case [26].

9.2.3.5 Clock Ring Configuration

The clock ring approach shown in Figure 9.48 combines aspects of the clock trunk, quadrant, and the single large buffer approaches to achieve a combination of moderately low skew, and moderately low delay without the possibly high routing-dependent skew of the pure single-buffer scheme. The ring approach also simplifies overall clock and signal routing conflicts in a two-layer metal process. The external clock is buffered at entry with a moderate- to large-scale buffer, which drives a clock ring that follows the die perimeter. The ring is not a widened trunk because typically less than 50 other buffers are driven off the ring, not the entire clock net. Therefore, with relatively low C_{load} on the ring, it is not widened. The extra capacitance of a widened ring would be relatively high (on a square die, the total ring length will be four times the length of a corresponding central trunk) and only increase skew and delay. The ring drives secondary buffers sized to drive balanced groups of flip-flops in the core of the chip. The aim of the ring on a large die is not to create a wholly isochronic perimeter, although this could be approached by driving the ring at multiple symmetric locations. Rather, the ring establishes a relatively low-skew reference perimeter from which any interior clock load can be reached either by a purely

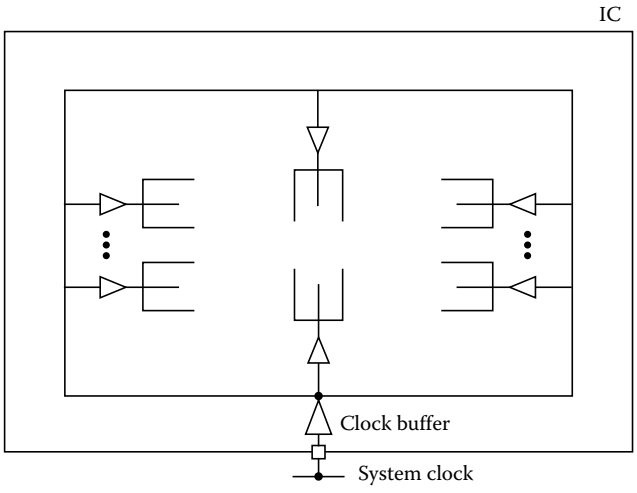


FIGURE 9.48 The clock ring concept.

vertical path or purely horizontal path (i.e., in a direct line of a single metal layer) no longer than one half the die size. The worst-case routing distance from the ring is the same as in the clock trunk scheme, but two thirds of all locations are within half that distance in the ring, whereas only one-half of all uniformly distributed loads are within half of the maximum distance of a central trunk. In practice, with good load balance on the secondary drivers, clock skew of 0.8–1.0 ns has been obtained in a designs of up to 30 K gates. If the secondary drivers are well balanced, skew in this architecture will depend primarily on the R_0C_0 delay from the ring driving point to its far side, around the periphery, typically 0.8 ns for a die of 350–400 mil. Relatively low chip delay is obtained by using the clock signal on the ring as an ''early clock'' with which to time I/O latches. The ring is electrically closed as this helps distribute the subbuffer capacitance and equalize delays, especially if driven at two opposing points.

9.2.3.6 H-Trees

The H-Tree is an area-efficient regular structure most suited to clock distribution in systems in which the synchronized modules are identical in size and placed in regular array. Figure 9.49 illustrates a 256 module H-tree tested on a 4 in. wafer by Keezer and Jain [17]. The scheme balances the R_0C_0 delay through the clock network by geometric symmetry so that the delay is nominally constant from the root to any leaf node. Loads are clocked only at leaf nodes of the tree. The minimum feature size of the process can be assumed to set the line width of the H-tree at its leaf nodes and each preceding level has progressively wider lines to maintain constant current density and to minimize impedance mismatch effects (at wafer scale and above) when no branching buffers exist. The H-tree is driven by a buffer at its root and may or may not have additional buffers at branching points. ASIC manufacturers have been able to achieve skew below 500 ps at fanouts of >5000 with experimental H-tree layouts in third-layer metal [22].

The H-tree approach is most practical only if an entire layer in a multilayer PCB or a third or fourth metallization layer in CMOS can be dedicated for H-tree clock distribution. Otherwise, the H-tree may encounter (or cause) a large number of routing blockages or require poly links which will disrupt

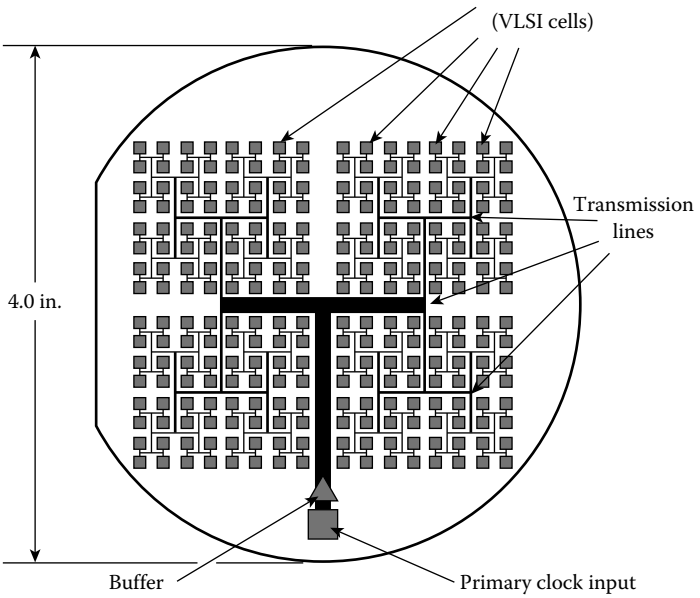


FIGURE 9.49 The H-tree concept illustrated in the form of a 256-cell passive H-tree for wafer-scale integration. (From Keezer, D.C. and Jain, V.K., *IEEE Int. Conf. WSI*, 1992, 168–175.)

the H-tree performance. In addition, many VLSI designs include memory cells or other hardcoded cells that are incompatible with the ideal symmetry of the H-tree. However, if a suitable layer is available for H-tree layout, it may be applied to random-logic designs by considering each leaf of the node as a clock supply point for all clocked loads within an isochronic region around each leaf node. The whole die is then tiled with overlapping isochronic regions, each fed from out of the plane by a leaf of the overlying H-tree. Each leaf of the H-tree might also use a variably sized buffer to drive the particular number of clocked loads of the random logic layout that fall in its particular leaf node zone.

Kung and Gal-Ezer [18] have given an expression for the time constant of an H-tree, which characterizes the total delay from root node to leaf:

$$\tau_H = 1.43\ N^3\left(3 - \frac{2}{N}\right)R_0C_0 \tag{9.6}$$

where an $N \times N$ array of leaf nodes is drives. Absolute delay rises as N^3 for large N. This in itself does not limit the clocking speed because, at least theoretically, more than one clock pulse could be propagating toward the leaf nodes within the H-tree. As a practical matter, however, the delay of the H-tree expressed in Equation 9.6 is essentially a rise time effect on the clock waveform. A slow rising edge out of the H-tree can lead to significant skew due to threshold variations in the leaf node buffers. These considerations apply to the on-chip context in which the H-tree clock network is dominated by RC diffusion delay effects. Equation 9.6 also describes an unbuffered H-tree. By placing buffers at selected levels of the overall tree, total propagation delay through the tree will increase, but the bandwidth of the tree may be preserved by effectively reducing N to the portions of the overall tree between buffers, in Equation 9.6. In contrast to the rapid bandwidth fall-off on-chip, at the multi-chip module level of system integration an H-tree may be designed from an impedance-controlled transmission line standpoint to obtain very high clock bandwidth. In experiments of this type, Bakoglu [3] has achieved 30 ps of skew at 2 GHz with a 16 leaf H-tree covering a 15×15 cm wafer area.

In an H-tree the total clock path length doubles each time one moves up two levels from the leaf nodes toward the root. Based on this, Kugelmass and Steiglitz [19] have shown that given σ_b and σ_w as the standard deviation of buffer delay (if present) and wire delay, respectively, the total delay of an H-tree considering buffers and wires has variance:

$$\sigma^2 = \sigma_b^2 \log_2(N) = \sigma_w^2 2\left(\sqrt{N} - 1\right) \tag{9.7}$$

and that the average case skew between any two leaf nodes is bounded by

$$E[\text{skew}] = \sigma_w 4\sqrt{\sqrt{(N-1)}\ln(N)} \tag{9.8}$$

where N is large and wire length effects dominate. Average case (or expected) skew is the maximum difference between clock times, averaged over many design trials, not the average clock time difference in any one design. Kugelmass and Steiglitz [19] also give results for the probability that a sample value of the skew exceeds the mean skew by a given factor in either an H-tree or a binary tree, based on assumptions that all wire length and buffer delay variables are independent and identically distributed:

$$P(\text{skew} > E[\text{skew}] + a) \leq \left[1 + \left(\frac{a}{E[\text{skew}]}\right)^2 48\ (\ln N)^2/\pi^2\right]^{-1} \tag{9.9}$$

where a is the amount of time by which the mean skew is exceeded. These expressions may be used to estimate skew-limited production yield at a given target clock speed.

TABLE 9.11 Comparative Performance of Clock Distribution Networks (8 × 8 Array of Loads Clocked at 31 MHz and Constant Total Power, 650 mW)

Scheme	Delay (ns)	Skew (ns)	Rise/Fall Time (ns)
3-level symmetric buffer tree	7	3	12.5
Single buffer H-tree	15	∼0.0	38
Clock trunk with branch buffers	13	4	14.2
Clock trunk with ganged branch buffers	14.2	2	16
4-pin-quadrant scheme, 2 buffers per quadrant	4.3	1.3	9

Source: Nigam, N. and Keezer, D.C., A comparative study of clock distribution approaches for WSI, *Proc. IEEE* 1993 *Int. Conf. WSI*, 1993, 243–251.

9.2.3.7 Delay, Skew, and Rise Time Comparison

The five clock distribution schemes described thus far were studied in a unified, experimental way by Nigam and Keezer [24] using HSPICE simulations. They compared each scheme on a 5 in. wafer holding an 8 × 8 grid of modules to be clocked. Each module presented at a total load of 2 pF and the interconnect R and C values were taken for a typical 2 μm double-metal CMOS process. Clock distribution lines were 10 μm wide for the buffer tree and the H-tree, except for its trunk, which was 40 μm. The clock trunk schemes used a 20 μm trunk width. All interconnect was modeled as distributed RC, with transmission line delay effects included. The results are tabulated in Table 9.11 and give an excellent overview of the relative characteristics of each method. The H-tree has essentially no skew, but has the highest delay and slowest clock edge, which can translate into skew due to threshold variations in the loads. The clock trunk has good skew and moderate delay. The best overall performance is achieved by the four-quadrant scheme, essentially by virtue of reducing the clocking area to one fourth of the overall size of the other clock networks.

9.2.3.8 Balanced Binary Trees

A balanced binary tree (BBT) is an unbuffered clock distribution tree in which each branch node has exactly two subtending nodes, and the delay from the root to all leaf nodes is made constant by placing branch points at the "balance point" of the two subtending trees from any node. BBTs are not simply clock buffer trees with fanouts of two. The significance of the BBT is that constant delay is achieved through multiple levels, without any buffers, and the BBT can be constructed by a fairly simple algorithm. Passive BBTs also may be used in practice to implement delay-equalized fanout networks between the active buffer levels of a larger buffered clock tree. The BBT concept should not be confused with the buffered clock tree concept in general, however. The key is that the generalized buffer clock tree does not have path delay equivalence if its buffers are removed, whereas the BBT has this property.

The basic ideas and methods of generalized balanced tree synthesis are explained in [6]. The clock tree that results has two branches at every node. Clocked loads appear only at the leaves of the tree. The line lengths at each level can be different than those at other levels of the tree, and the two line segments that are children of any node also can be of unequal lengths. The key, however, is that at each branch the total distance from the branch point to any subtending leaf via one outgoing path is equal to that in the other outgoing direction.

Figure 9.50 illustrates the basic procedure for BBT synthesis. The process works from the bottom up, by considering all leaf node positions, i.e., clock entry points for modules or isochronic local regions. Leaf nodes are subjected to a generalized matching or pairing process in which each node is paired with the other node closest to it in the Manhattan street length sense within the available routing channels. A first-level branch point is then defined at the midpoint on each line joining the paired leaf nodes. A similar pairing of the first-level branch points then defines a new set of line segments, each of which has two leaf nodes symmetrically balanced at its ends. Second-level branch are then defined at the RC balance point

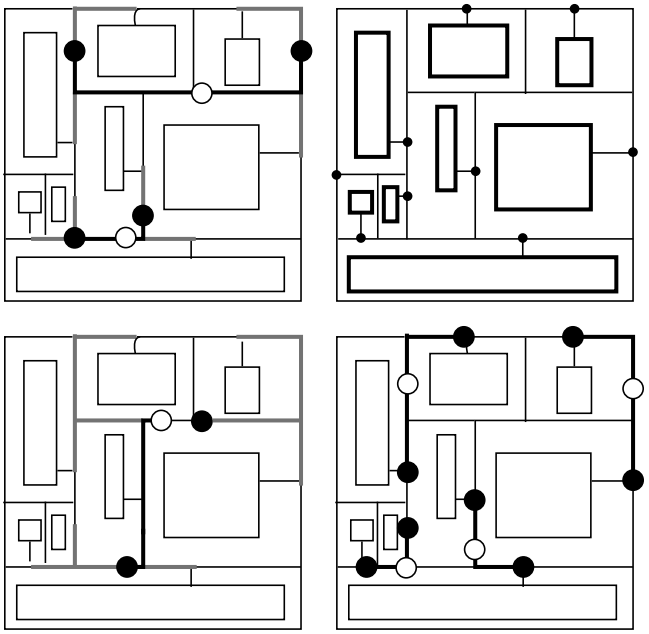


FIGURE 9.50 BBT synthesis in an eight-terminal net. Solid dots are roots of subtrees in the previous level; hollow dots are roots of new subtrees computed at the current level. (From Cong, J., Kahng, A., and Robins, G., *Proc. 4th IEEE Int. ASIC Conf.*, Sep. 1991, 14.5.1–14.5.4.)

on each line joining first-level branch points, and so on. Each iteration consists of finding the minimum path length matching of pairs of branch points from the previous iteration and defining the next set of branch points at the time-constant balance points on the line between just-matched, lower-level branch points. Assuming leaf nodes present equal loads, the first-level branch points are the midpoints of the pair-matching line segments. After that, the balance point for level i is the point on the line segment joining matched level $(i-1)$ branch points at which the total RC wiring time constants (simply total length if all wires and layers are identical) to the leaf nodes in the right and left directions are equal. This is repeated until only two branch points remain to be matched. A line is routed between them and driven by the clock signal at the final balance point, which defines the root of the BBT.

Clock trees developed in this way are in one sense highly structured and symmetric in that the total delay from root to any leaf is nominally constant, like the H-tree. Unlike the H-tree, however, the actual layout is irregular, allowing the BBT to accommodate the actual placement of cells and modules and to cope with the limited and irregular routing channels available in designs that do not use a completely regular layout.

Skew in BBTs has been considered theoretically and experimentally. Kugelmass and Steiglitz [19] showed that in a BBT with independent variations in the delay at each stage of the tree, with σ_0^2 variance, the expected skew is fairly tightly bounded by

$$E[\text{skew}] \leq \frac{4\sigma_0}{\sqrt{2\ln 2}} \ln N \tag{9.10}$$

where N is the number of leaf nodes of the tree.

Using the previous expressions we can compare the H-tree to a BBT. The comparison shows that when the regular structure of an H-tree is feasible, it is of relative merit for large fanouts because the expected skew grows more slowly ($O(N^{1/4}(\ln N)^{1/2})$) than the BBT tree in which expected skew grows as $O(\ln N)$.

For comparison, assuming 10,000 leaf nodes and the same σ_w per unit wiring length, the expected skew of the H-tree is about one half that expected of the BBT. This outcome is primarily because the BBT must be deeper (have more levels) than the H-tree for the same fanout.

Experimentally, Cong et al. [6] produced a large sample trials of 16- and 32-node BBT clock trees, synthesized on a 1000×1000 grid. It was shown that the BBT resulted in less than 2% of the skew from a corresponding minimum spanning tree (MST) for clock distribution to the same loads, even though the BBT had 24% to 77% more total routing length than the MST. The MST benchmark characterizes the skew that would typically result if the clock was routed as an ordinary signal net, with no special concern about skew.

An example of balanced clock tree synthesis supported by a gate array provider is [22], in combination with a three-level clock buffer tree hierarchy. Skew of $<$500 ps is achieved in 0.5 μm designs of up to 13,440 clocked loads. By using appropriately sized buffers and wire width at each level of the balanced tree, clock rise time is typically 0.8–0.9 ns at the terminal nodes. The clock tree compiler is invoked after floorplanning. The compiler takes into account the resistance and capacitance of different wire types, the length and width of wires, and the input capacitance of clock pins and buffers. Up to three active buffering levels can be used, with fanouts of up to 64, 14, and 15, respectively, from buffers at each level. The fanout subnet driven by each buffer is laid out as a (passive) BBT, so that the leaves of one balanced tree are the buffers that act as the roots of further passive binary balanced subtrees driven by the next buffering level. At the lowest level, local buffers each drive up to 15 loads via a final four-stage passive BBT.

9.2.3.9 Clocking Schemes Involving Phase-Locked Loops

A PLL is a negative-feedback control system in which the phase (and, implicitly, frequency) of a voltage-controlled oscillator (VCO) or phase shifter is brought into alignment, or to a predefined static phase offset, with respect to the phase of a periodic reference signal. The application of PLLs is most often to control skew and clock delay problems primarily at the multichip and interboard levels of system design.

Figure 9.51a shows how a PLL can be used to lock the on-chip clock phase at a selected point on-chip, to an external phase reference. Figure 9.51b shows the mid-trunk phase on a single-ended clock trunk being made to match the external phase reference. Here, the feedback line from the middle of the trunk to the PLL input is assumed to have negligible delay in itself because it is a metal line with only one standard load. Similarly, in the double-ended clock trunk scheme, the sense line can be connected one fourth of the way along the clock trunk. This will lock the internal system clock to the reference timing at two points on the clock trunk, as shown in Figure 9.51c, reducing overall skew to one fourth of that in the single-ended clock trunk scheme. The phase-sense line needs to be connected to the clock trunk at only one point because, by symmetry, the corresponding point from the other driver is similarly phase locked.

In general, when the phase-sense line has negligible delay, the clock phase at the sense point is driven into lock with the reference phase. Thus, in generating clock signals we can null out the delays of large buffers or drivers in the output circuits as well as their process and temperature-dependent variations, and, in general, coordinate clock and data phases at the inputs to another chip at any remote point by bringing the phase-sense line back from the actual point where the phase-controlled relationship is desired. In this way even the delay of an off-chip driver can be canceled out by including it within the PLL feedback loop. If delay in the phase-sensing feedback path is not negligible, then its effect is to advance the phase at the desired control point. The feedback delay can be compensated by a matching delay in the forward path from the VCO to the phase-sensing point, or at the PLL input. An inverter in the feedback signal path is also a convenient way to cause a 180° phase shift between referenced and VCO without requiring any loop delay.

With the addition of frequency divider (divide by N) in the feedback path, as in Figure 9.51d, the VCO operates at N times the frequency of the reference clock input. For $N = 2^n$ frequency multiplication, the feedback divider can be a simple ripple counter of n toggle flip-flop stages. For other multipliers, a synchronous counter is usually used. The delay-matching element at the PLL phase detector input compensates for delay from the feedback path divider. On-chip frequency multiplication can ease a

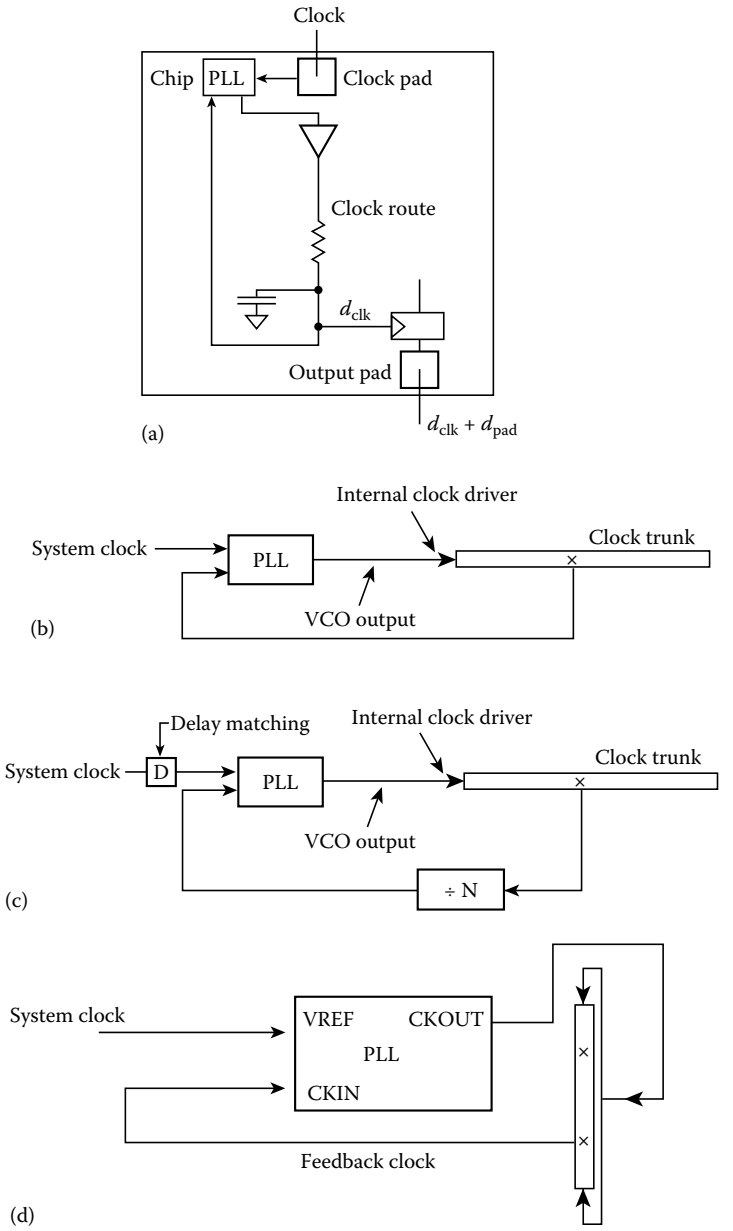


FIGURE 9.51 PLLs for skew and delay control: (a) canceling internal and clock net delay. (From Weste, N. and Eshraghian, K., *Principles of CMOS VLSI Design*, Addison-Wesley, Reading, MA, 1993, 335.) (b) Halving the skew in a single-ended clock trunk, (c) reducing double-ended clock trunk skew to one-fourth of the single-ended trunk, (d) one-chip frequency multiplication. (From LSI Logic Corp., *Phase-Locked Loop Application Note*, LSI Logic Application Note, Nov. 1991.)

number of system-level design problems. The overall system clock rate need not be equal to that of the fastest chip in the system. Transmission line effects across the relatively long distances of PCBs or backplanes can be reduced by operating at a lower clock frequency outside of the system ICs. The lower frequency of system reference distribution may also reduce power, and usually assists in meeting radiated

emission specifications for electronic equipment. Because a PLL regenerates the clock in each IC, a considerable amount of clock edge slew rate control can be used on the external system clock, further easing EMI and power supply switching noise problems. The difficulty of retaining clock waveform integrity getting on- and off-chip at high frequencies through inductive packaging and bonding leads is also eased for the same reason.

9.2.3.10 PLLs for CMOS

A block diagram of a PLL is shown in Figure 9.52a. The VCO exhibits a positive monotonic frequency of oscillation in response to a control voltage, characterized by the slope of its frequency vs. voltage curve. The loop filter, $H(f)$, is of a general low-pass characteristic, often of an all-poles design to avoid any jitter peaking (or AC gain) in the closed loop transfer function of the PLL. The loop filter must provide a DC coupled path between the phase detector and VCO. The phase (and/or frequency) detector compares the VCO output phase to the input reference phase and generates an output signal that is either of a DC nature or has a DC component that is proportional to the phase difference between the reference and feedback signal. The phase detector is characterized by the rate of change in the DC component of its output vs. phase input difference in V/rad.

A phase detector that is commonly used because of its all-digital nature and suitability for CMOS integration is Gardner's phase-frequency detector (PFD) [11] with charge pump outputs. The PFD produces an output that goes toward V_{DD} or V_{SS} in the presence of a negative or positive frequency offset, respectively, thereby slewing the VCO toward the lock frequency. Once in frequency lock, the PFD

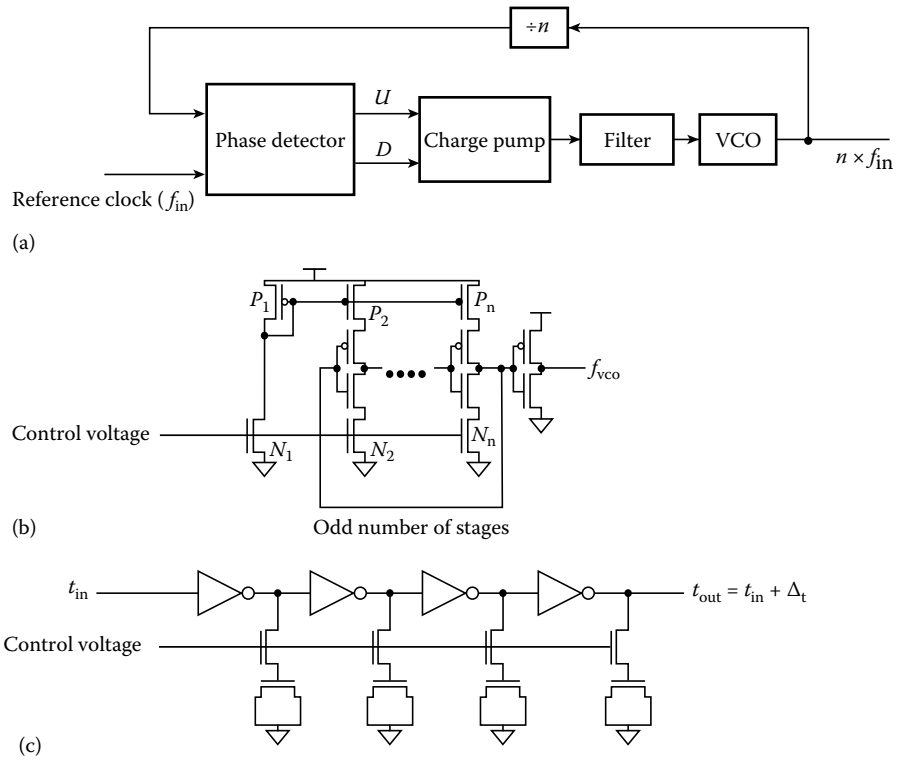


FIGURE 9.52 CMOS PLL circuits: (a) basic PLL block diagram, (b) CMOS VCO based on current-starved inverters, and (c) VCDL. (From Weste, N. and Eshraghian, K., *Principles of CMOS VLSI Design*, Addison-Wesley, Reading, MA, 1993, 336.)

produces pump up/pump down signals that vary in proportion to the time difference between reference and feedback clock edges at the PFD input. These pulse-width modulated signals drive a charge pump with a tristate buffer arrangement to either hold, bleed off, or supply charge to a capacitive storage element (i.e., the loop filter), thereby adjusting (and filtering) the voltage on the VCO control node to minimize the phase difference at the phase detector inputs.

For CMOS clocking system applications, the VCO is usually a form of a stable multivibrator in which the switching speed dependence of a CMOS inverter on its *n*-transistor pulldown current is exploited, as shown in Figure 9.52b. The VCO control voltage regulates the current flow in, and hence the speed of, each inverter stage through the extra *n*-transistor stage added to each inverter. Any odd number of such stages connected in a loop will oscillate, but now the relaxation period is voltage controlled. Weste and Eshragian [31] describe a 13-stage ''current-starved inverter'' VCO based on this approach. A related VCO design is based on varying the load capacitance seen by each inverter (in a chain of inverters) by applying the VCO control voltage to an *n*-MOSFET in series with the gate of another transistor configured as a capacitive load [16]. An on-chip RC loop filter can also be constructed from a CMOS transmission gate biased as a resistor and MOS gates used as capacitors (source and drain both connected to V_{SS}) [31].

If frequency multiplication is desired, a PLL with a true VCO is required. Otherwise, many PLL applications can use a voltage-controlled delay line (VCDL) in conjunction with a ''raw-clock'' input signal, as in Figure 9.52c. All the phase-lock feedback principles are the same, except we phase-shift the raw-clock input as required, rather than controlling the VCO oscillation phase. This eliminates the risk of the PLL ever failing to lock-in and may be simpler to fabricate. On the other hand, system design must take into account a more limited range of phase-shifting ability (a full half-cycle of phase control range may require a lot of delay stages) and to make sure initial delays are nominally centered within the positive-only delay control range of the VCDL. Another more subtle point is that while a VCO introduces a perfect integrator ($1/s$ term) in the PLL closed loop response, a VCDL does not. A VCDL, therefore, should not be simply substituted for a VCO without revisiting the closed loop response characteristics for noise bandwidth and possible jitter peaking.

Special power, grounding, and testing considerations apply when a PLL is used. A PLL is basically a linear circuit, so noise is especially important. Particularly when a frequency multiplying PLL is used, the VCO power supply should be well decoupled from system noise, and the input phase reference should be highly stable, as the PLL output clock will have N times the reference's phase noise. Noise voltages coupled into the analog PFD output and LPF signal path are similarly converted into phase noise that is N times worse than in a $\times 1$ PLL. Leadless on-chip decoupling capacitors are recommended as are dedicated power and ground pins for the PLL. The R and C components for the loop filter are often off-chip. In this case it is important that they are connected (depending on the $H(s)$ configuration) to the same analog ground reference as the VCO and PFD. The VCO output, or a divided down version of it, should be brought to an external pin for lock-in validation and as an aid in possible global system clock tuning. For testability, several other separate pins are typically required for independent access to I/O of the PFD, LPF, and VCO each. An on-chip PLL can require up to six or more pinouts.

9.2.3.11 Anceau's PLL Scheme

Anceau [1] developed a PLL-based approach for large systems in which modules are well-defined, relatively independent, and could be entirely self-timed if not for the need to avoid metastability in communication with other modules. Anceau recognized two natural system scales which are isochronic below different maximum frequencies. One is a global region encompassing the entire system, with a clock period determined by propagation distance delays, or, on-chip, by RC diffusion delays. The isochronic rate for this scale defines a slower clock rate for a system-wide communication bus. The second type of clocking region is smaller local regions which can run at full speed and are characterized by critical logic path delays and lumped capacitive loads within modules, not distance-dependent delays. Each smaller region will be free to operate in an almost self-timed mode.

The clocking style within each module (e.g., logic type and number of clock phases) can be as appropriate for the individual modules. Skew at the highest clock speeds in the system need be considered only within each module, except that timing must be controlled when reading the common data bus to avoid metastability. This is done by reference to the active edge of the slower-rate communications clock (comm_clk), formed by dividing down the master module clock frequency. The rising edge of comm_clk strobes the enabled driver data onto the bus. All other nondriving modules in the comm_clk cycle read the bus on an internal clock edge that is kept away in time form this transition in comm_clk, for metastability avoidance. Figure 9.53 illustrates the overall scheme. A PLL phase locks the module clock at a predefined angle relative to the comm_clk, thus keeping the raw module clock away from the transition times in the lower rate communication clock. The read timing is then safe because it is always preceded by the comm_clk transition on which new date were strobed to the bus. A monostable triggered by the comm_clk edge can be used inside each module as a delay generator to prohibit any bus read in the metastable region. This way, as long as modules write to the bus only on the comm_clk edge, other modules that read the bus will never do so at a moment when the bus data are still in transition.

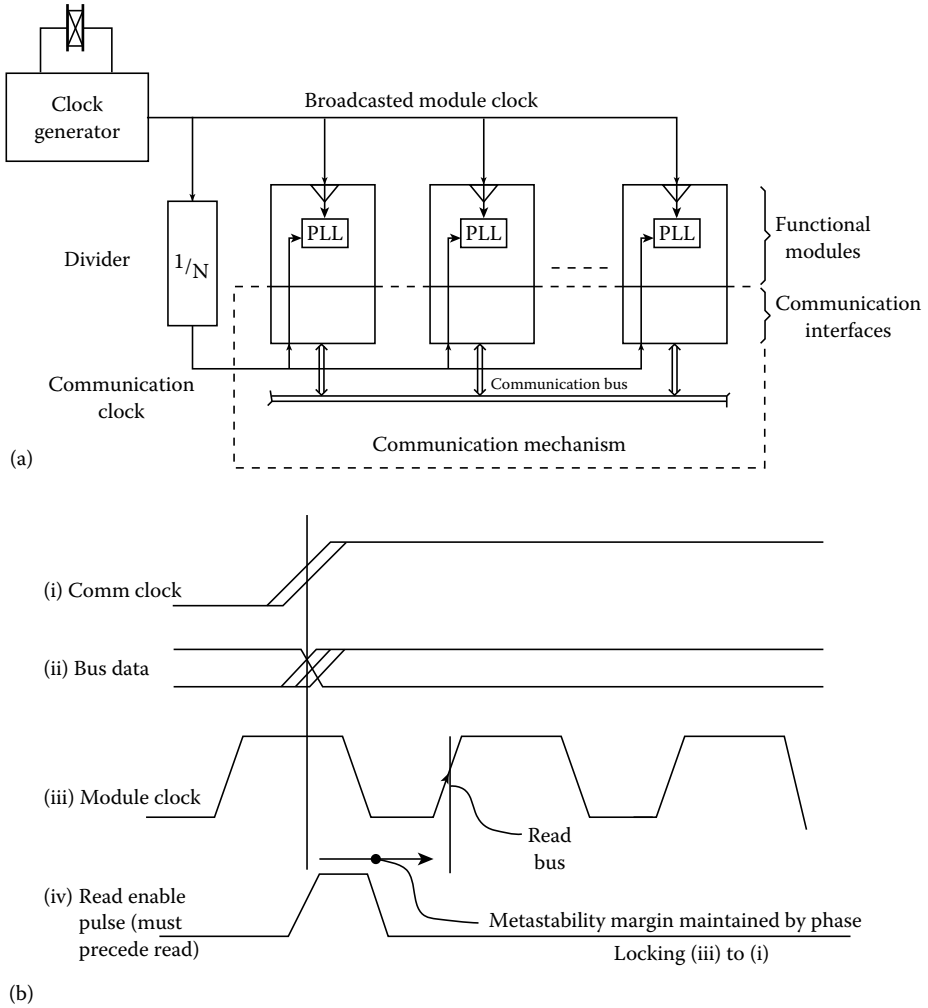


FIGURE 9.53 Anceau's scheme for metastability avoidance: (a) system architecture and (b) interface timing. (Adapted from Anceau, F., *IEEE J. Solid-State Circuits*, SC-17(1), 51, 1982.)

9.2.3.12 Grover's Interval-Halving PLL Scheme

A novel PLL-based approach to clock distribution in "electrically" large systems can synchronize all clocks in a large system on a single clock line [13,14]. In this scheme any number of nonisochronic points arbitrarily located on a single- or double-conductor reference line independently derive clock that is in absolute phase-lock to a common system-wide reference time. The central principle is that the time between appearances of an isolated pulse traveling down and back on a reference line is the same regardless of the point of observation, as shown in Figure 9.54a. This figure plots the trajectory in space-time of an isolated pulse that travels from a site at one end of a line and is returned at the end of the line to its origin (where it is electrically terminated). Figure 9.54a is drawn for the most general case of physically separate go and return conductors, looped at the right-hand end ($x = D$), but the space-time trajectory of the isolated pulse is identical if the line is a single conductor open-circuited at the end and driven by an impedance-matched source. Equivalently, a tristate buffer can terminate and regenerate the returning pulse at the end of the line for on-chip use. In either case, it is evident that the instant in time that is halfway between the outgoing and returning pulse edges is the same for all points of observation on the line, regardless of the propagation velocity of the line, i.e.,

$$\frac{t_1[x] + t_2[x]}{2} = t_1[D] = t_2[D] \equiv t_{\text{ref}} \tag{9.11}$$

where

$t_1[x]$, $t_2[x]$, are the times when the traveling pulse edge passes position x in Figure 9.54a
$t[D]$ is the time the reference pulse edge reaches (and departs) the reflection point

This time, called t_{ref}, is the midpoint between the two pluses as seen at every point of observation on the line.

This principle is adapted for single-line, skew-compensated clock distribution by periodic injection of a reference pulse onto a single conductor, reflection of this pulse at the end of the reference line, and generation of a local clock at all stations, such clock being phase-locked to the interval mid-time by a special interval halving PLL (IHPLL) circuit, as outlined in Figure 9.54b. The phase detector in the IHPLL is considerably simpler than the conventional PFD used in many CMOS PLL designs. This method can be adapted easily to a two-line operation, in which a full duty cycle waveform, rather than a narrow pulse, can be used to drive the looped reference line path. The reference line can be looped at one end and driven at the other, or split into two terminated lines, routed together as a pair, and driven together as shown in Figure 9.54c. In the latter case, all modules lock to the reference edge arrival at mid-path, rather than at its end.

In either of the two-conductor configurations, an edge-triggered, set–reset flip-flop function (Figure 9.54d) is the required phase detector. The two-conductor IHPLL approach avoids the need for an end-reflection or a tristate returning line driver, but requires layout of the reference line so that distances from the end are the same on both directions of the path at every tapping point. This is not hard to achieve on-chip, as the two halves of the looped path could be laid out identically. At the system level, however, the single-line variant has the advantage that uncontrolled cable and tracking lengths can be used without concern about delay equivalence in the return path, and an absolute minimum of cabling, connectors, and tracking is required for clock distribution. The interconnect is the same as that for a system in which clock is directly wired to all modules with a single line. However, this would normally be possible only if the whole system were one isochronic region.

Grover [13] reports experimental skew under 1 ns over 30 m on a coaxial cable which has an uncompensated delay of 147 ns. It was also shown that in the presence of the effects of the transmission line on the traveling reference pulse, the linear component of switching time error on the traveling reference pulse contributes no skew to the resultant clock phase. A phase-shifter variant of this scheme

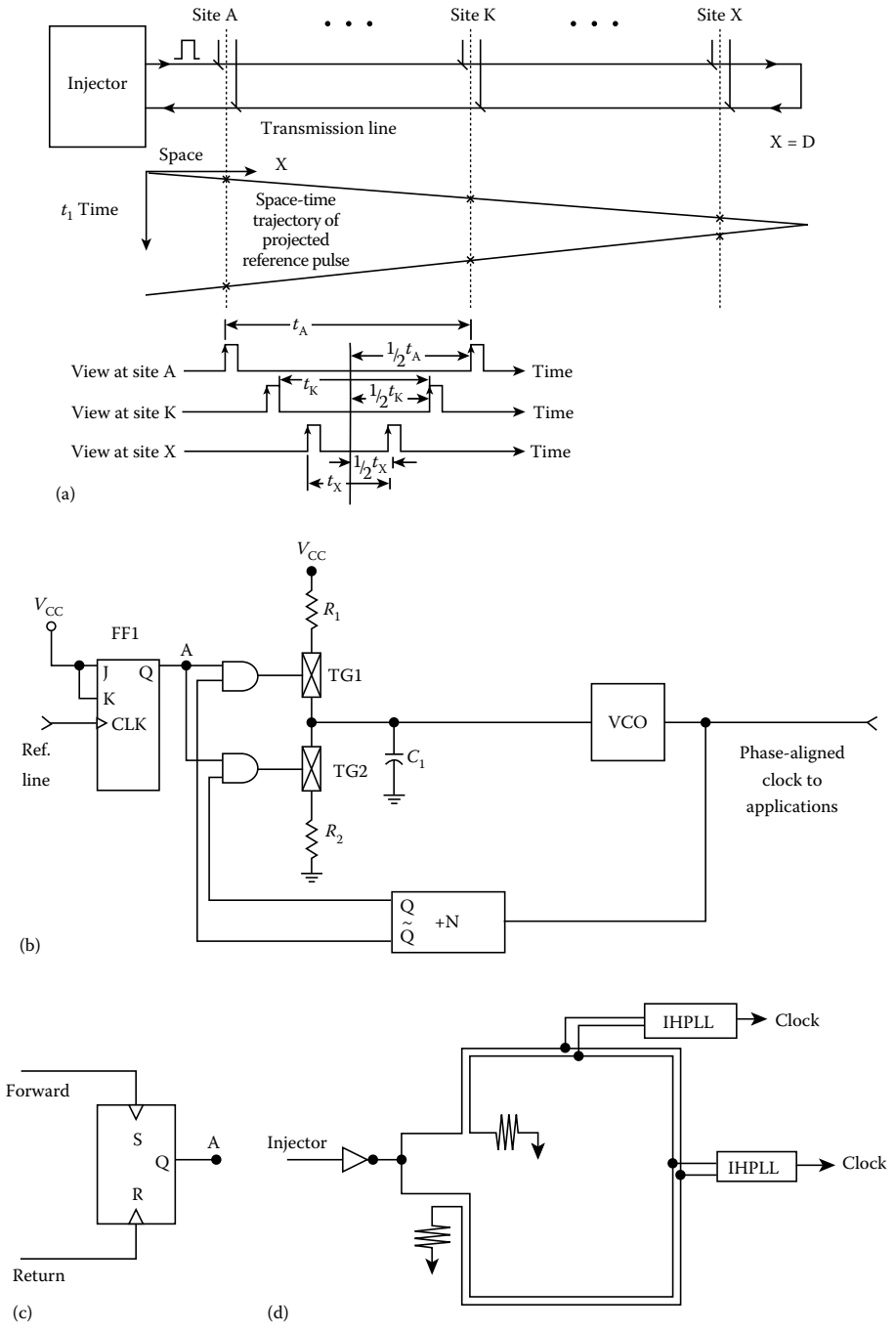


FIGURE 9.54 Grover's interval-halving clock distribution scheme: (a) interval-halving reference-time principle, (b) IHPLL circuit for single-line skew-compensated clock distribution, (From Grover, W.D., Method and apparatus for clock distribution and distributed clock synchronization, US Patent #5,361,277, Issued Nov. 1, 1994.) (c) phase detector for 2-line operation, and (d) driving two lines rather than one looped line.

uses a separate line to distribute raw-clock, which is then adaptively phase shifted at each point into the low-skew global phase by a voltage-controlled delay under the same feedback control sensing arrangement. Two-line operation, phase shifter, and other variations are described further in [14].

With this scheme, hierarchical clock distribution networks with delay-controlled cabling, delay-tuning, and numerous temperature- and load-dependent intermediate buffers may be replaced by one conductor with arbitrary routing. Both EMI and conducted noise are reduced by buffer driver elimination and because of the reduced average power of the reference pulse compared to the full clock signal. It may also be possible to add new clock-deriving taps in service, offering a growth path that is not limited by a predesigned clock–tree fanout limit. In many applications hybrids of reduced-depth clock trees, fanning out from skew-compensated roots on a single-line clock system of this type, may give the best combination of techniques.

Anceau's and Grover's schemes are similar in that a reference line is distributed to all modules and a PLL generates a local clock at each module. However, in the Anceau scheme modules do not run phase synchronously. Actual skew between modules remains arbitrarily high at the module clock frequency because the comm_clk line is set slow enough to be isochronic over the whole system. Actual delay in comm_clk, which is significant at the higher module rate, is not compensated at modules. Each module derives only enough information to coordinate its bus accesses with other modules, at the slower comm_clk frequency. In Grover's scheme, however, truly synchronous full-speed global clocking of all modules is achieved by returning the signal on the reference line (by reflection or looping) and exploiting the interval-halving time-reference principle and IH-PLL to cancel global skew. Gate-to-gate interaction on any clock cycle is feasible between modules in this case, as opposed to interaction only through the metastable-avoidance bus interface protocol.

9.2.3.13 Clock Tuning in Large Systems

In large systems, clock tuning at the chip, circuit pack, shelf, and rack levels of physical equipment may be required for the highest performance. Circuits to permit clock tuning can be a tapped delay-line circuit with a programmable selector, or (on a circuits card) a printed-in set of loops to be shunted out as needed by a suitcase jack, or a voltage-controlled varactor clock delay buffer. All of these are described in [15]. In general, to aid in the tuning process, one pin on each IC should be devoted to give external observability of the worst-case (if known) clock phase form inside the IC. This way the tuning process can compensate for the delays through I/O pads and clock buffers in large ICs.

Tuning begins by designing cable lengths, tracking, and connectors so that clock paths have nominally equivalent delays. The active tuning process then measures and adjusts relative delay starting from the master clock source to predefined levels of tuning points (TPs) electrically farther from the master clock source, denoted by TP_0. The delay measurement and adjustment repeats through lower level tuning points until the clock on in every IC is tuned. In going from the first to successive tuning points, it is preferable to refer delay measurements directly back to TP_0 each time. This may, however, be physically unmanageable, in which case delay tuning to level TP_n can be relative to TP_{n-1} although overall error relative to TP_0 will be higher in the relative tuning scheme. For systems that must grow in service, operational (i.e., in-service) signals should be the basis of the delay measurement, not requiring off-line signals or patterns.

One convenient way of indirectly measuring delay between points that are not easily accessed simultaneously for oscilloscope measurements is to make an oscillator out of the signal path to be measured. If the number of inversions in the path between TPs is odd, then an MUX can be switched to loop the tuning point signal at TP_n back to the TP_{n-1} driving point. A frequency counter can then measure the oscillation frequency, providing data to support automatic clock delay adjustments at the subordinate tuning point.

One mainframe computer used an automatic tuning scheme in which a clock phase-shifter chip produced multiple, slightly time-shifted copies of the clock on each system PCB. Individually selected delayed clock instances were then supplied to each IC on the board through a programmable crosspoint matrix IC. Each clock-receiving IC also provided several internal clock observation outputs to support

delay measurements down to the gate level. After automatically measuring the delays of the observable internal clocks in up to 30 ICs per board, the on-board clock selection matrix was programmed, giving each IC its best clock phase for overall system timing margins and minimal skew [30].

9.2.4 Future Directions

9.2.4.1 Current-Steered Logic

One way of reducing power supply noise injection from clocking is with current-steered logic. Experiments on differential current mode flip-flops in CMOS predict very short setup times (300–500 ps). Such devices would be very quiet electrically and much less susceptible to varying load effects on delay than on conventional CMOS. On the other hand, such devices might be about twice as large as conventional CMOS flip-flops and require more power. Using the bi-CMOS ability to integrate ECL type structures with CMOS may, however, be part of the solution for clocking very high speed medium-scale integrated devices.

9.2.4.2 Reduced Voltage Swing

Another potential method for reducing clocked load power consumption is to reduce voltage swings. Some experiments indicate a significant reduction in clock-related power and ground noise, but skew and delay objectives are more difficult to meet as the devices slow down in response to lower switching voltage swings.

9.2.4.3 Mixed Technology

Here, clock speed increases are envisaged by using current mode logic circuits selectively to implement critical timing paths in otherwise all-CMOS systems. New power reductions may also be obtained with ECL-based, high-speed, serial-multiplexed interfaces to replace wide buses which have many parallel CMOS drivers.

9.2.4.4 $\overline{Q}$ Elimination

Most logic families provide flip-flops with both Q and $\overline{Q}$ outputs as standard cells. An approach that could potentially have clock-related switching current and power is to eliminate the $\overline{Q}$ output buffer and develop corresponding logic synthesis tools to utilize inverted inputs and other logic means to assemble logic function without the $\overline{Q}$ outputs form flops. In one experiment of which the author is aware, $\overline{Q}$ buffers were removed, halving overall flip-flop power consumption, at the cost of only a 5% degradation in clock-to-Q delay.

9.2.4.5 Dedicated Layer for Clock Distribution

A number of workers are advocating or are already using dedicated third-layer metal for clock distribution. This affects process cost, but the advantages can be significant in high-performance applications. Third-layer metal is lower than other layers in resistance and capacitance. By moving the clock net, which is the largest single net in many designs, out of the other layers, routability of all other signals is improved and floorplanning simplified. Moreover, the clock tree can avoid uncertain delays due to unpredictable routing or due to polysilicon links in series when routing in fewer shared signal layers. In addition, noise due to clocking can be more easily isolated in the third-layer metal approach.

9.2.4.6 Optoelectronic Clock Distribution

Optical clock distribution takes advantage of the three-dimensional nature of imaging optics to remove all but the last-stage buffering levels of the clock distribution tree from the plane of the circuit, thereby eliminating multiple stages of buffering and metallization for clock routing. Figure 9.55 is an overview of the basic idea proposed by Clymer and Goodman [7]. The optical clock signal is generated off-chip and drives a laser diode at the top of the figure. The optical beam is expanded onto a transmission hologram,

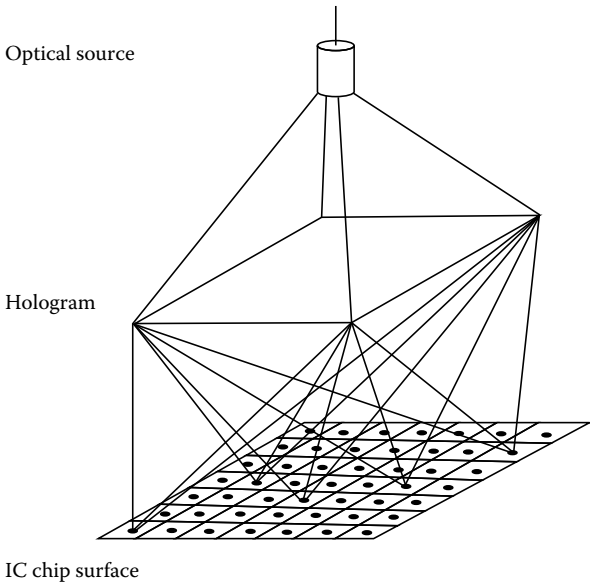


FIGURE 9.55 Clymer and Goodman's concept for wafer-scale optical clock distribution. (From Clymer, B.D. and Goodman, J.W., *Opt. Engin.*, 25(10), 1103, 1986.)

which focuses the light intensity onto predefined locations where optical detectors are fabricated into the wafer or die. The optical signal is detected, amplified, and used to drive a local clock generator-buffer which supplies a local isochronous regions. The optical path length differences are not equalized in this scheme because the optical path velocity is so high as to make the all optical path delays negligible as compared with the diffusion and lumped capacitive delays that determine the clock rare of the electronic system. When sources, detector, and packaging for this type of approach are developed, the potential exists for very low skew—high-speed clock distribution, with greater on-chip densities by eliminating most clock routing. One of the main challenges is in attaining uniform response times from the optical detector–amplifier combination (which tend to be sensitive to feature size variations) and the development of sources in the optical wavelength range for photodetectors that can be fabricated within the conventional CMOS circuit environment.

9.2.4.7 Reconfigurable Clock Nets

In WSI systems, where a single short in very large clock net may disable an entire wafer-level system, Fried [10] advocates methods of restructuring a clock net to enhance yield, primarily through the addition of a controllable tristate output stage to clock buffers within the clock distribution network. This way failed portions of the clock net can be isolated, or, with redundant interconnect and buffers, they may clocked by an alternate path. In particular, tristate buffers may be programmed on or off to select clock for each module from redundant connections to the central clock net, or to simply isolate failed clock net subregions from the drivers of unfailed portions.

References

1. F. Anceau, A synchronous approach for clocking VLSI systems, *IEEE J. Solid-State Circuits*, SC-17(1), 51–56, Feb. 1982.
2. M. Afghahi and C. Svensson, A unified single-phase clocking scheme for VLSI systems, *IEEE J. Solid-State Circuits*, 25(1), 225–233, Feb. 1990.

3. H. B. Bakoglu, *Circuits, Interconnections and Packaging for VLSI*. Reading, MA: Addison-Wesley, 1990, chap. 8.
4. S. Boon, S. Butler, R. Byrne, B. Setering, M. Casalanda, and A. Scherf, High performance clock distribution for CMOS ASICS, *Proc. IEEE 1989 Custom Integrated Circuits Conf*., San Diego, CA, pp. 15.4.1–15.4.5, 1989.
5. M. A. Cirit, Clock skew elimination in CMOS VLSI, *Proc. IEEE Int. Symp. Circuits Syst*., New Orleans, LA, pp. 861–864, 1990.
6. J. Cong, A. Kahng, and G. Robins, On clock routing for general cell layouts, *Proc. 4th IEEE Int. ASIC Conf*., Rochester, NY, pp. 14.5.1–14.5.4, Sep. 1991.
7. B. D. Clymer and J. W. Goodman, Optical clock distribution to silicon chips, *Opt. Engin*., 25(10), 1103–1108, Oct. 1986.
8. R. H. Cramback, C. M. Lee, and H. S. Law, High-speed compact circuits with CMOS, *IEEE J. Solid-State Circuits*, SC-17, 614–619, June 1982.
9. D. Dobberpuhl et al., A 200 MHz 64-b CMOS microprocessor, *IEEE JSSC*, 27(11), 1555–1567, Nov. 1992.
10. J. Fried, Power and clock distribution for WSI systems, *Proc. IFIP Workshop on Wafer Scale Integration*, G. Saucier and J. Trilhe, Eds. Amsterdam: North-Holland, pp. 127–141, 1986.
11. F. M. Gardner, Charge-pump phase-locked loops, *IEEE Trans. Commun*., COM-28, 1849–1858, Nov. 1980.
12. L. A. Glasser and D. W. Dobberpuhl, *The Design and Analysis of VLSI Circuits*. Reading, MA: Addison-Wesley, 1985, chap. 6.
13. W. D. Grover, A new method for clock distribution, *IEEE Trans. Circuits Syst. Part I*, 41(2), 149–160, Feb. 1994.
14. W. D. Grover, Method and apparatus for clock distribution and distributed clock synchronization, US Patent #5,361,277, Issued Nov. 1, 1994.
15. M. W. Johnson and M. Graham, *High-Speed Digital Design: A Handbook of Black Magic*. Englewood Cliffs, NJ: Prentice Hall, 1993, chap. 11.
16. M. G. Johnson and E. L. Hudson, A variable delay line PLL of CPU-coprocessor synchronization, *IEEE J. Solid-State Circuits*, 23(5), 1218–1223, Oct. 1998.
17. D. C. Keezer and V. K. Jain, Design and evaluation of wafer scale clock distribution. *IEEE Int. Conf. WSI*, San Francisco, CA, pp. 168–175, 1992.
18. S.-Y. Kung and R. J. Gal-Ezer, Synchronous vs. asynchronous computation in very large scale integrated (VLSI) array processors, *Proc. SPIE*, 341, 53–64, 1982.
19. S. D. Kugelmass and K. Steiglitz, An upper bound on expected clock skew in synchronous systems, *IEEE Trans. Comput*., 39(12), 1475–1477, Dec. 1990.
20. LSI Logic Corp., *Phase-Locked Loop Application Note*, LSI Logic Application Note, Nov. 1991.
21. LSI Logic Corp., *Clock Scheme for One Micron Technologies*, Rev. 1.1, LSI Logic Application Note, Aug. 1992.
22. LSI Logic Corp, *Clock Distribution Schemes for 300K Technologies*, Rel. 2.0, LSI Logic Application Note, May 1993.
23. A. M. Moshen and C. A. Mead, Delay-time optimization for driving and sensing of signals on high-capacitance paths of VLSI systems, *IEEE Trans. Electron Devices*, ED-26, 540–548, 1979.
24. N. Nigam and D. C. Keezer, A comparative study of clock distribution approaches for WSI, *Proc. IEEE 1993 Int. Conf. WSI*, San Francisco, CA, pp. 243–251, 1993.
25. W. A. Samaras, The CPU clock system in the VAX 8800 family, *Digital Tech. J*., 4, 34–40, Feb. 1987.
26. T. Saigo, S. Watanabe, Y. Ichikawa, S. Takayama, T. Umetsu, K. Mima, T. Yamamoto, J. Santos, and J. Buurma, Clock skew reduction approach for standard cell, *Proc. IEEE* 1990 *Custom Integrated Circuits Conf*., Boston, MA, pp. 16.4.1–16.4.4, 1990.
27. C. L. Seitz, System timing, in *Introduction to VLSI Systems*, C. Mead and L. Conway, Eds. Reading, MA: Addition-Wesley, 1980, chap. 7.

28. J. Shyu, A. Sangiovanni-Vincentelli, J. Fishburn, and A. Dunlop, Optimization-based transistor sizing, *IEEE J. Solid-State Circuits*, 23(2), 400–409, Apr. 1988.
29. D. Tanksalvala et al., A 90 MHz RISC CPU designed for sustained performance, *IEEE Solid-State Circuits Conf.*, San Francisco, CA, pp. 52–53, Feb. 1982.
30. K. D. Wagner, A survey of clock distribution techniques in high speed computer systems, Report CRC 86-20. Stanford, CA: Stanford University Center for Reliable Computing, Dec. 1986.
31. N. Weste and K. Eshragian, *Principles of CMOS VLSI Design: A Systems Perspective*, 2nd edn. Reading, MA: Addison-Wesley, 1993, pp. 317–335 (clocking strategies), pp. 334–336 (PLL methods), pp. 685–689.
32. D. F. Wann and M. A. Franklin, Asynchronous and clocked control structures for VLSI based interconnection networks, *IEEE Trans. Comput.*, C-32(3), 284–293, Mar. 1983.

9.3 MOS Storage Circuits

Josephine C. Chang and Bing J. Shev

In a large digital system, a sequence of operations must be performed for a particular function. The results of each operation depend on the results of previous operations. Therefore, the outputs of a logic circuit block typically depend not only on present input signals, but also on the history of the inputs. A CL circuit becomes more useful if it is combined with memory elements. To construct a sequential system, the most common and straightforward way is to employ a central clock to synchronize the sequence of operations.

Instead of using memory elements in a sequential system, we can use dynamic logic circuits to store temporary data. With the building blocks of inverters and transmission gates, the MOS transistors can be used as dynamic storage components to store data temporarily on the device capacitances. Dynamic storage is widely used in MOS technologies because of the simplicity of the required circuitry. Because a memory element such as a static circuit latch occupies a large area and consumes power, elimination of latches has a positive effect on circuit density and power consumption. However, the disadvantages of dynamic logic gates include high transient power disturbances and less noise margins in some applications [1].

Dynamic logic circuits design is based on the synchronized movement of charge through the MOS circuit. A typical capacitance value associated with a logic gate is on the order of a few femtofarads, with means the amount of charge $Q = CV$ dynamically stored on the capacitance is on the order of femtocoulombs. Therefore, perturbations from ideal behavior can become critical to the operation of a circuit.

9.3.1 Dynamic Charge Storage

The MOS technologies have two attractive features that lead to an efficient way to store data momentarily. These two features are the extremely high input impedance of MOS transistor and the ability of a MOS transistor to function as a nearly ideal electrical switch. In order to store the charge on a capacitive node, the node must be isolated from both the power supply and ground. Various types of storage nodes can be realized in CMOS technologies. For example, charge can be stored at a node between sources (or drains) of two MOS transistors such as nMOS–nMOS, pMOS–nMOS, and nMOS–pMOS [2]; or the source (or the drain) terminal of one MOS transistor connected to the gate terminal of a second MOS transistor. Because the stored charge will leak away over time, this circuit is termed ''dynamic storage circuit.''

Figure 9.56 shows the schematic diagrams of three combinations of source–drain connection. The distinction among the three connection types comes from the difference in voltage transmission levels for nMOS and pMOS gates.

Dynamic charge storage requires clocking the data at a sufficiently high rate so that the charge on the various nodes does not lead away significantly. Typically, this requires a minimum refresh rate of 500 Hz to 1 kHz, corresponding to a charge storage time of about to 2 ms.

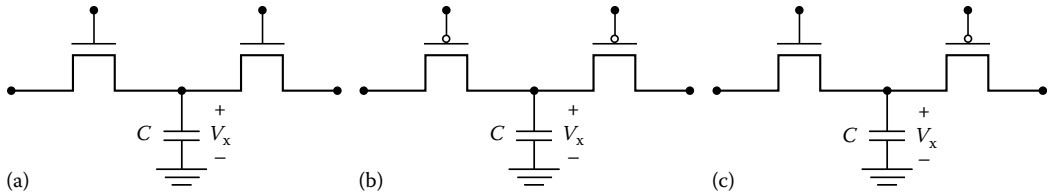


FIGURE 9.56 MOSFET source–drain connection storage nodes. (a) nMOS–nAMOS; (b) pMOS–pMOS; and (c) nMOS–pMOS.

9.3.1.1 nMOS–nMOS

An nMOS transistor is perfect for transmitting logic 0 signals, but imperfect for transmitting logic 1 signals due to the threshold voltage loss through the transistor. The voltage level of V_x which can be stored on the capacitor C is therefore limited by

$$0 \leq V_x \leq (V_{DD} - V_{th,n}) = V_{max} \tag{9.12}$$

where

$$V_{th,n} = V_{th\ 0,n} + \gamma_n\left(\sqrt{2\phi_{Fn} + V_{max}} - \sqrt{2\phi_{Fn}}\right) \tag{9.13}$$

Charge storage on an nMOS–nMOS node is affected by the leakage paths through the p-type bulk to the ground. This affects the long-term storage of a logic 1 value.

9.3.1.2 pMOS–pMOS

A pMOS–pMOS node is the complement storage component of an nMOS–nMOS node. The voltage level of V_x is limited by

$$V_{min} = \left|V_{th,p}\right| \leq V_x \leq V_{DD} \tag{9.14}$$

where

$$V_{th,p} = V_{th\ 0,p} + \left|\gamma_p\right|\left(\sqrt{\left|2\phi_{Fp}\right| + (V_{DD} - V_{min})} - \sqrt{2\left|\phi_{Fp}\right|}\right) \tag{9.15}$$

Because both p-channel MOS transistors have n-type bulks which are connected to V_{DD}, this type of storage node receives leakage current from the power supply. The logic 1 values can be held indefinitely, but the logic 0 values can only exist for a limited period of time.

9.3.1.3 nMOS–pMOS

A complementary nMOS–pMOS storage node can benefit from both nMOS and pMOS in transmitting logic 0 and logic 1, respectively. The voltage level which is stored on the capacitance is in the range of

$$0 \leq V_x \leq V_{DD} \tag{9.16}$$

If the maximum input value of V_{DD} is transmitted through the pMOS transistor and the minimum input value of 0 V is transmitted through the nMOS transistor. On the other hand, if the case is reversed, the

maximum input value is entered through the nMOS transistor and the minimum input value is entered through the pMOS transistor, then the voltage range is reduced to

$$\left|V_{\text{th,p}}\right| \leq V_{\text{x}} \leq V_{\text{DD}} - V_{\text{th,n}} \tag{9.17}$$

This type of operation should be avoided because it greatly reduces the noise margins. In a standard nMOS–pMOS storage node both leakage paths to the power supply and ground exist. The ability to retain logic 0 and logic 1 values depends on which leakage path dominates.

9.3.1.4 Source–Gate Connection

This type of storage node is the connecting point between the source terminal of a pass transistor and a gate terminal of another MOS transistor [3]. Electrical charge can be temporarily stored on or removed from the gate terminal of the second transistor. When the gate terminal of the pass transistor is at a logic low value, the pass transistor is turned off, and the charge on the gate terminal is isolated. This charge determines the stored logic value. If the stored charge is perfectly isolated, the logic value would be stored indefinitely. In a practical situation, the isolation is less than perfect, primarily because of leakage through the reverse-biased diode operation between the source diffusion region of the pass transistor and the substrate. In addition, leakage also can occur through the pass transistor. With the continuous advances in VLSI technologies, subthreshold leakage through the channel of the pass transistor becomes more important due to scale-down in device sizes. Leakage currents alter the node voltage, which may lead to a logic error.

Two major problems arise in maintaining the integrity of a stored logic state. First is the parasitic conduction paths in the transistors that lead to charge leakage. Leakage currents alter the node voltage, which may cause a logic error. The second problem is charge sharing, which occurs when two isolated storage nodes become connected by a switching event and must equalize their voltages by redistributing charge. Charge sharing may result in a logic error, or may block logic propagation entirely.

9.3.1.5 Charge Sharing

Beside charge leakage, a problem called charge sharing may also damage the integrity of a stored logic state. Charge sharing occurs when a dynamic charge-storage node is used to drive another isolated node in a switching network [4]. Typically, when two capacitors with different voltages are connected by a pass transistors, as shown on Figure 9.57, charge sharing may occur. When the pass transistor is turned on, the voltages on the capacitors equilibrate to some intermediate value. In Figure 9.57, capacitors C_1 and C_2 are in parallel when the transmission gate is conducting. This forces the voltages across C_1 and C_2 to be equal. If the two capacitors are charged to different initial voltages, charge sharing will occur when the

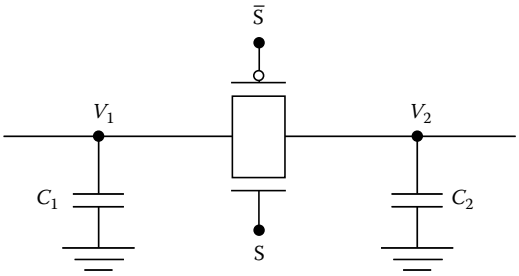


FIGURE 9.57 Charge-sharing-prone structure.

transmission gate turns on. Let the initial voltage charge on C_1 be V_1 and Q_1, and the initial voltage and charge on C_2 be V_2 and Q_2. The initial charge balance equation is

$$Q_1 + Q_2 = C_1V_1 + C_2V_2 \tag{9.18}$$

After the transmission gate turns on, the final charges on C_1 and C_2 become Q_1' and Q_2', respectively, and both capacitors are charged to the same value V'. The final charge balance equation is

$$Q_1' + Q_2' = (C_1 + C_2)V' \tag{9.19}$$

By applying the charge-conservation principle, we can obtain

$$V' = \frac{C_1V_1 + C_2V_2}{C_1 + C_2} \tag{9.20}$$

and

$$Q_1' = \frac{C_1}{C_1 + C_2}(C_1V_1 + C_2V_2) \tag{9.21}$$

A precharged circuit might work incorrectly due to charge-sharing errors, which could occur inside the pulldown network or at the output circuit. To control a precharged circuit, a gated clock can be present only at the input of the bottom transistor, while all other inputs to the gates of transistors in series in the pulldown chain must have a stable signal over the same clock phase to prevent charge-sharing problems. A ''sneak path'' is created when two pass transistors in series are both turned on at the same time and one is connected to V_{DD} while the other is connected to the ground. Charge can leak through this sneak path.

9.3.2 Shift Register

A frequent use of dynamic storage circuits is the shift register. Shift registers are most often used to provide temporary storage of digital signals. The shift register storage can be used as a simple way to delay the arrival of a signal for a specific number of clock cycles. Shift register storage is also frequently used as the temporary memory for a sequential logic circuit. In general, shift registers provide dense, limited access memory for many applications in digital integrated circuits.

9.3.2.1 Simple Shift Register

Figure 9.58 is the schematic diagram of a multistage MOS shift register, with each stage composed of a pass transistor and an inverter [5]. The nonoverlapping clock waveforms Φ_1 and Φ_2 are used. Assume that a logic signal is placed at the input of the first shift register stage while the Φ_1 clock is low and the transmission gate of the first stage is turned off. Next, when the Φ_1 clock goes high, if the signal at the input to the first stage is held constant, it will be propagated to the input of the inverter in the first stage. After a short delay, the output of the first inverter will provide the inverted logic signal to the input of the second shift register stage. At this time, the Φ_2 clock is low and the transmission gate in the second stage will not pass this signal. When the clock values change so that Φ_2 becomes high, the transmission gate of the second stage will propagate the output signal of the first stage to the second inverter, and then the output of the second stage is produced. This signal will be stopped by the transmission gate of the third stage because Φ_1 is low while Φ_2 is high. This sequence continues through the shift register chain as the clock signals alternate, causing the input signal to propagate through the shift register stages. The data are

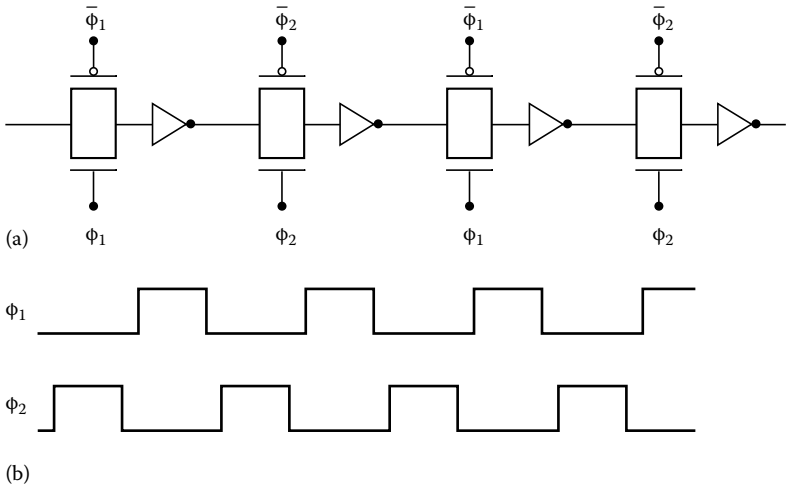


FIGURE 9.58 (a) A four-stage MOS register and (b) nonoverlapping waveforms of ϕ_1 and ϕ_2.

stored on the capacitances associated with the gate terminals of the inverter. The transmission gate acts as the switch that lets charge flow into and out of the capacitors when they are turned on. The charge is trapped at the capacitor when the transmission gate is turned off.

Each time the Φ_1 clock changes to a high value, the shift register input signal will propagate to the gate of the first inverter and the output signal of the first stage will be produced. A sequence of alternating Φ_1 and Φ_2 clock signals will cause an input signal to propagate through the whole structure at the rate of two stages of the shift register for each complete cycle of the clock signals. After N clock cycles, a logic input value will have shifted through $2N$ stages of the shift register chain. When a two-phase clock is used to control a shift register, it is important that the two clock phases do not overlap. If both phases of the clock were high simultaneously, a data value could propagate through multiple stages during the clock overlap time. This would cause erroneous operation of the shift register.

9.3.2.2 Parallel Shift Register

Several copies of the multistage shift registers can be combined in parallel with the same clock lines to form a parallel shift register to transmit a group of signals in lock-step fashion. Such a parallel shift of 8, 16, or 32 data bits is often used in microprocessor circuits. The basic structure of this set of shift registers demonstrates two principles which are important for the efficient geometrical layout of digital circuits. The data for the shift register flow from left to right while the control signals (Φ_1 and Φ_2 clocks) flow from top to bottom. Such an orthogonal structure of data paths and control signals within a circuit module is widely used to provide a regular organization of logic circuits within a VLSI chip. Physical layout of the shift register stages can be mirrored with respect to the ground and V_{DD} lines. This mirroring technique allows shared power and ground connections and reduces required circuit layout area. It is important to minimize the size of the basic shift register stage because this stage is repeated many times in a large shift register.

9.3.2.3 Clocked Barrel Shifter

A "barrel shifter" is a wraparound shifter that forms a very useful switch array [6]. The basic layout is shown in Figure 9.59. The inputs are labeled I_i; the shift controls $\Phi_2 \cdot \mathrm{SH}_i$, and the outputs O_i. The input lines run horizontally while the output lines run vertically. The operation of the first shift register

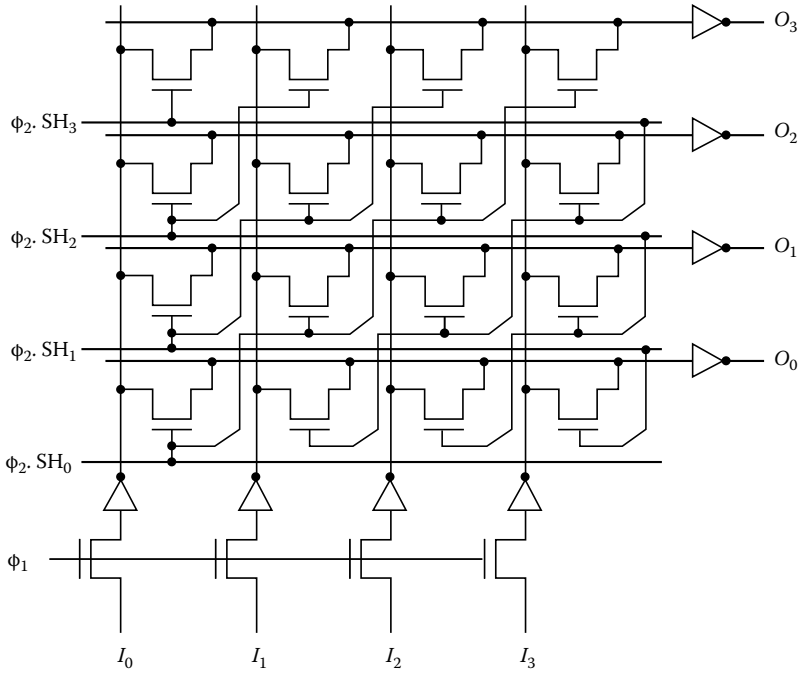


FIGURE 9.59 A four-bit clocked barrel shifter.

stage is the same as explained earlier. In the second stage for four output signals from the four inverters in the first stage can be shifted without changing the order or each signal can move up one, two, or three locations.

9.3.3 Dynamic CMOS Logic

The dynamic CMOS logic design consists of dynamic circuits based on precharging the output node to a particular level when the clock is at the logic 0 level. During the precharge phase, the inputs to the circuits change. When the clock is at the logic 1 value, the output of the logic gate may be pulled to a complementary value, depending on the input conditions.

The choice of using static or dynamic logic is dependent on many criteria. When low-power performance is desired, it appears that dynamic logic has some inherent advantages in a number of areas including reduced switching activity due to hazards, elimination of short-circuit dissipation, and reduced parasitic node capacitances. Static logic circuits have advantages on charge sharing and precharge operation.

Static circuits design can exhibit spurious transitions due to races. These spurious transitions dissipate extra power over that required to perform the computation. The number of these extra transitions is a function of input patterns, internal state assignment in the logic design, delay skew, and logic depth. Although it is possible with careful logic design to eliminate these transitions, dynamic logic intrinsically does not have this problem because any node can undergo at most one power-consuming transition per clock cycle [7].

Short-circuit currents are found in static CMOS circuits. However, by sizing transistors for equal rise and fall times, the short-circuit component of the total power dissipated can be kept to $<$20% of the dynamic switching component. Dynamic logic does not exhibit this problem, except for those cases in which static pullup devices are used to control charge sharing or when clock skew is significant. Dynamic

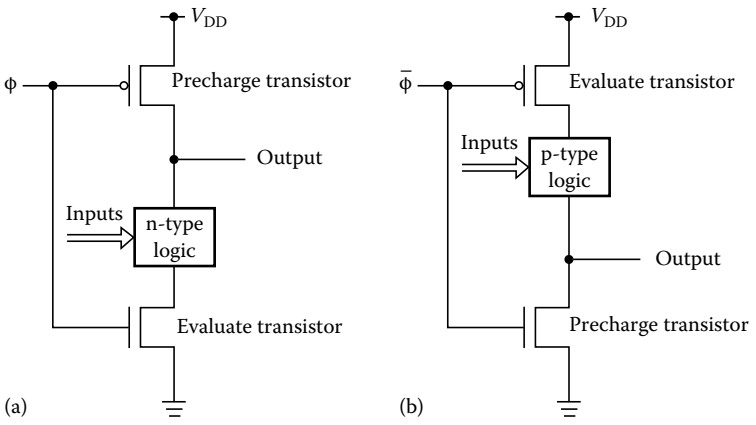


FIGURE 9.60 A precharge–evaluate logic gate.

logic typically used fewer transistors to implement a given logic function, which directly reduces the amount of capacitance being switched and thus has a direct impact in the power-delay product.

9.3.3.1 Precharge–Evaluate Logic

The schematic diagram of a basic precharge–evaluate logic is shown in Figure 9.60. It consists of an *n*MOS logic structure whose output node is precharged to V_{DD} by a pMOS precharge transistor; and conditionally discharged by the *n*-transistor network connected to the ground. Alternatively, an *n*-transistor precharge to the ground and a pMOS logic structure to conditionally discharge to V_{DD} may be used. A single-phase clock Φ is used for high-speed operation. For the former case, the precharge phase occurs when the clock Φ is low. The path to the ground is activated via the *n*-transistor network when the clock Φ is high. The input capacitance of this logic gate is the same as a pseudo-nMOS gate which has a single *p*-transistor, with the gate connected to the ground, as a load device. The pullup time is better than a pseudo-nMOS gate by virtue of the active switch but the pulldown time is increased due to the ground switch.

9.3.3.2 Clocked CMOS Logic

Clocked CMOS logic (C^2MOS) gates were originally used to build low-power dissipation logic gates. The reasons for the reduced dynamic power dissipation stem mainly from metal–gate CMOS layout considerations. The main use of such logic structures is to form clocked structures that incorporate latches or interface with other dynamic forms of logic structure. The gates have the same input capacitance as regular complementary gates, but larger rise and fall times due to the serially connected clocking transistors.

The schematic diagram of a clocked CMOS logic gate is shown in Figure 9.61. In this circuit, the clocked transistors are placed in series with the transistors in the *p*- and *n*-type logic blocks. The primary use of C^2MOS is in dynamic shift registers. In a C^2MOS dynamic shift register, the *p*-type logic black is a *p*-transistor network and the *n*-type logic is an *n*-transistor network. All transistors can normally be chosen as minimum-size devices because each stage is only required to drive the capacitance of an identical shift register stage.

Although the C^2MOS circuit requires the same number of transistors, external connections, and clock phases as the standard CMOS dynamic shift register, the layout is greatly simplified because the source/drain regions of the two *p*-channel transistors can be merged, and the corresponding regions of the two *n*-channel transistors can be merged. This feature helps to reduce circuit capacitance, number of contacts, and layout area.

Operation of the C^2MOS circuit is quite simple. The gates of the *p*MOS pullup transistor and the nMOS pulldown transistor of the inverter are both connected to the input signals. For a valid logic input,

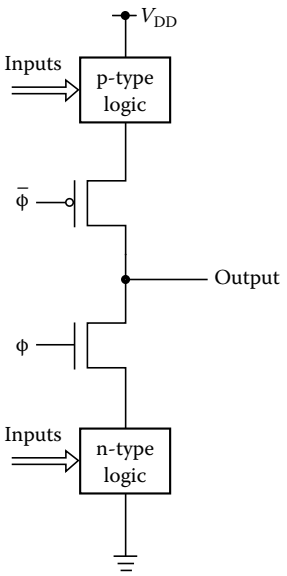


FIGURE 9.61 A clocked CMOS logic gate.

one of these transistors is turned off while the other is turned on. Clocked transistors placed in series with the pullup and pulldown transistors serve to connect these transistors to the output node when the clock Φ is high. If the input signals match the *n*-type logic portion, the output storage node will be discharged. Otherwise, the output storage node will be charged. When the clock Φ is low, the output node will remain in its present state. In contrast to other clocked logic circuits, the output of C^2MOS is available during the entire clock cycle, although it is actively driven only when the clock is high. A C^2MOS circuit is more susceptible to interference from the load circuit attached to the stage because the load capacitance is the storage node for the dynamic charge.

9.3.3.3 Domino CMOS Logic

The domino logic gate design can provide glitch-free cascades of nMOS logic structures. It is a modification to the clocked CMOS logic gate to allow a single clock to precharge and evaluate a cascaded set of dynamic logic blocks. A domino logic gate consists of two elements: a precharge–evaluate logic stage followed by a static inverter buffer at the output, as shown in Figure 9.62. The logic gate can be built in two forms: mostly *n*-transistors and mostly *p*-transistors. During the precharge phase when clock Φ is low, the output node of the dynamic gate is precharged high, and inverted by the static buffer to provide a logic 0 output for the domino CMOS gate. As subsequent logic stages are driven by this buffer, transistors in subsequent logic blocks will be turned off during the precharge phase. When the gate is evaluated, the node voltage of the logic stage is conditionally pulled down according to the input signal values. If the logic condition of the gate is satisfied, the node voltage is pulled down. It is inverted by the static buffer to provide a logic 1 output. Each gate in sequence can make at most one transition ($1 \rightarrow 0$). Hence, the buffer can only make a transition from ($0 \rightarrow 1$). In a cascaded set of logic blocks, each state evaluates and causes the next stage to evaluate in the same manner as a stack of dominos fall. Any number of logic stages may be cascaded, provided that the sequence can evaluate within the given clock phase. A single clock can be used to precharge and evaluate all logic gates within a block.

The structure has some limitations. First, only noninverting structures can be constructed. Second, each logic gate must be buffered. Finally, in common with a clocked CMOS gate, charge redistribution can be a problem. The effects of these problems can be minimized. For example, in complex logic circuits, such as ALUs, the necessary XOR gates may be implemented conventionally as complementary gates and driven by the last domino circuit. The buffer is often needed to drive large capacitive load and does not contribute to any extra cost.

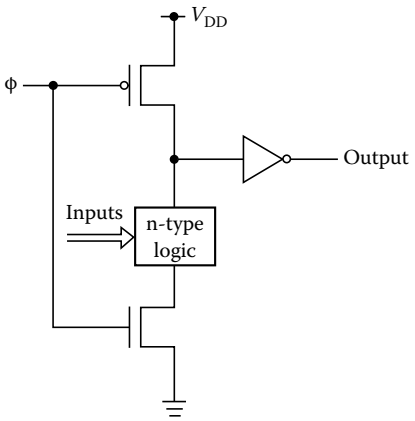


FIGURE 9.62 A domino CMOS logic gate.

Static storage of charge can be realized by a domino logic gate by including a weak *p*-transistor. A weak *p*-transistor is one that has low gain, which is realized with a small W/L ratio. It should have a small gain in order not to fight with the pulldown transistors, yet to balance the effects of leakage. This will allow low frequency or static operation when the clock is held high. In this case, the pullup speed could be an order of magnitude slower than the pulldown speed. Notice that the precharge transistor may be eliminated if the time between evaluation phases is long enough to allow the weak pullup to charge the output node.

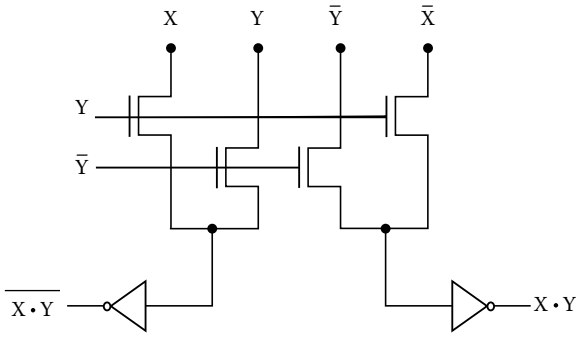


FIGURE 9.63 A CPL AND/NAND cell.

A domino logic gate has advantages over a simple precharge–evaluate logic structure. For example, the static buffer provides output-driving capability to either V_{DD} or the ground. In the precharge–evaluate logic gate the output can be favorably driven only to the ground in response to logical conditions, not to V_{DD}. When the logic condition of the precharge–evaluate gate is not satisfied, dynamic charge storage at the output must maintain the logic 1 value. The dynamic logic portion of a domino CMOS gate always has a fanout of 1, thereby simplifying device sizing within the gate structure.

9.3.3.4 Complementary Pass-Transistor Logic

The complementary pass-transistor logic (CPL) gate is constructed by using an nMOS pass-transistor network for logic function and eliminating the pMOS latch [8]. It consists of complementary inputs and outputs, nMOS pass transistor logic network, and CMOS output inverters. Figure 9.63 is the schematic diagram of a CPL AND/NAND cell. The pass-transistors function as pulldown and pullup devices. Thus, the pMOS latch can be eliminated, allowing the advantage of differential circuits to be fully utilized. One attractive feature of the CPL gate is that complementary outputs are produced by the simple four-transistor circuits. Because the logic 1 value level of the pass-transistor outputs is lower than the supply voltage V_{DD} by the threshold voltage of the pass-transistors, the signals must be amplified by the output inverters. In addition, the CMOS output inverters shift the logic threshold voltage and drive the capacitive load. The logic threshold shift is necessary because that of the output inverter is lower than half of the supply voltage, due to the lowering of the logic 1 value.

The CPL gate is attractive because fewer transistors are required to implement important functions. However, a CPL gate has two basis problems. First, the threshold drop across the single-channel pass-transistor results in reduced current drive and hence slower operation at a reduced supply voltage. This is important for low-power design because it is desirable to operate at the lowest possible voltage levels.

Second, because the logic 1 input value at the regenerative inverters is not V_{DD}, the pMOS device in the inverter is not fully turned off, and hence direct-path static power dissipation could be significant. To solve these problems, reduction of the threshold voltage has proven effective, although if taken too far it will incur a cost in dissipation due to subthreshold leakage and reduced noise margins.

9.3.3.5 Cascade Voltage Switch Logic

The cascade voltage switch logic (CVSL) gate is a differential style of logic circuit design requiring both true and complement signals to be routed to the gates [9]. Two complementary nMOS switch structures are constructed and then connected to a pair of cross-coupled p pullup transistors as shown in Figure 9.64a. When the inputs switch, node voltages Q and $\overline{Q}$ are either pulled up or down. Positive feedback applied to the p pullup transistors causes the gate to switch. The logic trees may be further minimized from the fully differential form using logic minimization algorithms. This version is slower than a

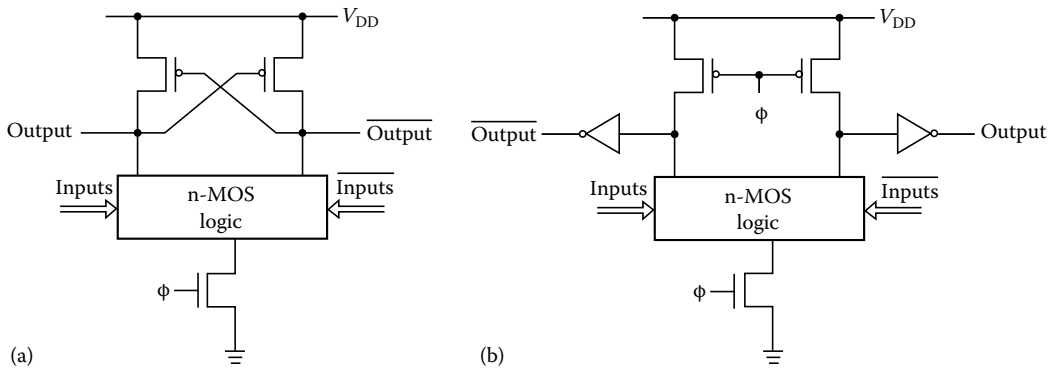


FIGURE 9.64 A CVSL gate: (a) static version and (b) dynamic version.

conventional complementary gate employing a *p*-tree and an *n*-tree because during the switching action, the *p* pullup transistors must compete with the *n* pulldown tree. The schematic diagram of a dynamic charge-storage version of the CVSL logic gate design is shown in Figure 9.64b. It consists of two domino logic gates with complementary input logic trees. The advantage of CVSL gate over a domino logic gate is the capability to generate a complete logic function rather than just the noninverting logic function. However, extra silicon area is needed.

9.3.3.6 NORA CMOS Dynamic Logic

NORA logic is capable of handling signal race problems in transmission pates [10]. It is based on dynamic CMOS logic, but uses latches instead of transmission gates to control signal flow. In a NORA logic dynamic nMOS and pMOS, logic circuits are cascaded into a C^2MOS latch. Figure 9.65 shows the schematic diagrams of both Φ stage and $\bar{\Phi}$ stage. Static inverters are provided at the outputs of dynamic circuits to realize logic inversion. This allows direct implementation of arbitrary functions without modification. In the Φ stage, the logic circuit used $\Phi = 0$ for precharge and $\bar{\Phi} = 1$ for evaluation. The latch accepts data when $\Phi = 1$ and holds the data when $\Phi = 0$. No new data can be accepted during the hold time. The operation in the $\bar{\Phi}$ stage is similar when reversing clock signals are used.

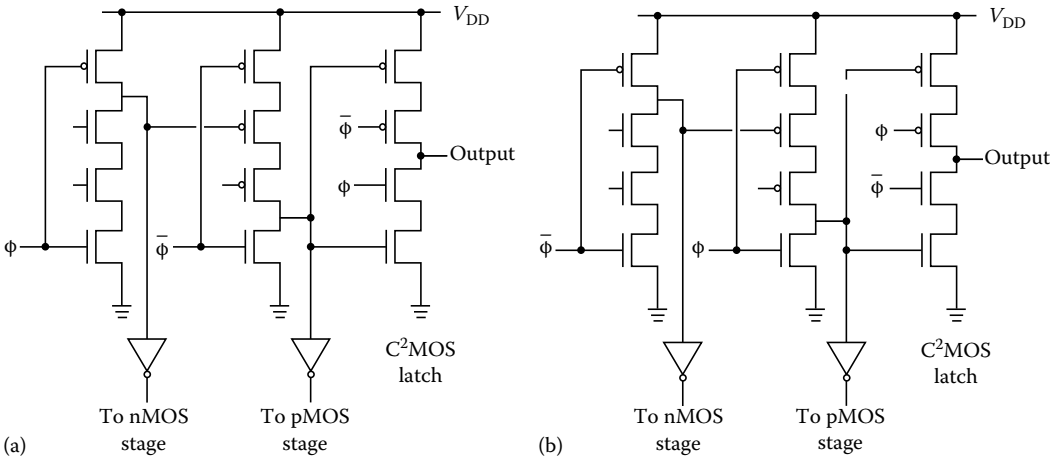


FIGURE 9.65 NORA clock stages: (a) Φ stage and (b) $\bar{\Phi}$ stage.

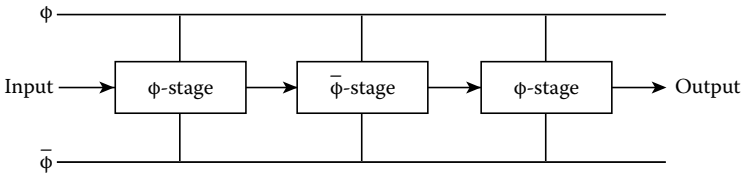


FIGURE 9.66 A NORA chain.

By alternating Φ and $\bar{\Phi}$ clock stages makes NORA chains well suited for pipelined logic. The schematic diagram of a generic structure of a NORA chain is shown in Figure 9.66 [2]. Logic flows through the chain at a rate set by the clock. The problem of logic races by using transmission gates as latches between logic circuits has been eliminated because of the dynamic C^2MOS latch circuit.

References

1. N. Wang, *Digital MOS Integrated Circuits*, Englewood Cliffs, NJ: Prentice Hall, 1989.
2. J. P. Uyemura, *Circuit Design for CMOS VLSI*, Boston, MA: Kluwer Academic, 1992.
3. R. L. Geiger, P. E. Allen, and N. R. Strader, *VLSI Design Techniques for Analog and Digital Circuits*, New York: McGraw-Hill, 1990.
4. A. Mukherjee, *Introduction to nMOS and CMOS VLSI Systems Design*, Englewood Cliffs, NJ: Prentice Hall, 1986.
5. L. A. Glasser and D. W. Dobberpuhl, *The Design and Analysis of VLSI Circuits*, Reading, MA: Addison-Wesley, 1985.
6. E. D. Fabricius, *Introduction to VLSI Design*, New York: McGraw-Hill, 1990.
7. A. Chandrakasan, S. Sheng, and R. W. Brodersen, Low-power CMOS digital design, 27(4), 473–484, April 1992.
8. K. Yano et al., A 3.8-ns CMOS 16 × 16-b multiplier using complementary pass-transistor logic, *IEEE J. Solid-State Circuits*, 25(2), 388–395, April 1990.
9. N. Weste and K. Eshraghian, *Principles of CMOS VLSI Design*, Reading, MA: Addison-Wesley, 1993.
10. N. F. Goncalves and H. J. De Man, NORA: A racefree dynamic CMOS technique for pipelined logic structures, *IEEE J. Solid-State Circuits*, 18(3), 261–266, June 1983.

9.4 Microprocessor-Based Design

Roland Priemer

9.4.1 Introduction

During the past three decades, microprocessors have become components that are routinely and widely used in machines and systems that engineers design. This is due to their flexibility and ability to perform tasks at a low cost. Because they are programmable, microprocessors are used to achieve operations of devices and systems with complexity that we have come to take for granted. Engineers are embedding microprocessors into systems that are being employed in virtually all fields of human endeavor.

The competition among the numerous manufacturers of microprocessors has brought about a great variety of microprocessors to increase their suitability in ever-widening fields, more products, and new markets. Moreover, turnaround times have become short enough and costs have come down enough so that system and application engineers have the option to specify the design of a customized microprocessor to meet their application-specific performance requirements.

This section is intended to introduce the reader to design with a microprocessor. It is assumed that the reader is acquainted with the material that is generally covered in an introductory course on digital systems. The goal is to help the designer who is not experienced with the design of microprocessor-based systems to come to a basic understanding of what is involved. Two representative and significantly different microprocessors will be used to do this. These are the Zilog Z80, an enduring general-purpose microprocessor, and the Motorola M68HC11, also called a microcontroller. These and similar microprocessors have been used in many diverse kinds of dedicated applications.

The design and development of the hardware and the software are two broad aspects to utilizing a microprocessor. Here, the emphasis is on digital hardware design. To understand the role that a microprocessor plays in the design process we must understand the operation of several major building blocks. These are memory, architecture of a microprocessor (programmer model), the system bus and timing of bus signals, and supportive devices to interface these building blocks together. At the level of design to be considered here, hardware and software are and should remain inseparable. From both points of view, we gain a better understanding of the other. However, due to space limitations, no significant presentation of programming in assembly language is included in the sequel. Instead, software is only utilized to see how it influences hardware requirements and how it makes hardware work.

9.4.2 Features of a Microprocessor-Based System

Conceptually, Figure 9.67 depicts a microprocessor-based system. The inputs to the microprocessor are all binary (can only take on one of two values) or digital signals. However, the origins and meanings of these signals can be very diverse. Inputs can be due to manually activated switch closures from buttons or keypads or they can be due to switches that are embedded within sensors that detect, for example, presence or absence of an object, pressure that is over or under a certain level, temperature that is above or below a particular temperature, presence or absence of an ultrasound or light beam, voltage that is above or below a desired value, etc. An input can come from a real-time clock. Inputs can also be information about the status or readiness of devices that receive outputs from the microprocessor-based system.

A single digital signal is called a bit, and it can only take on one of two logic values, i.e., logic one or logic zero, which in the circuits that we will use correspond to 5 or 0 V, respectively. Typically, these circuits are designed to accept as logic one any voltage in the range 3.5–5, and accept as logic zero any

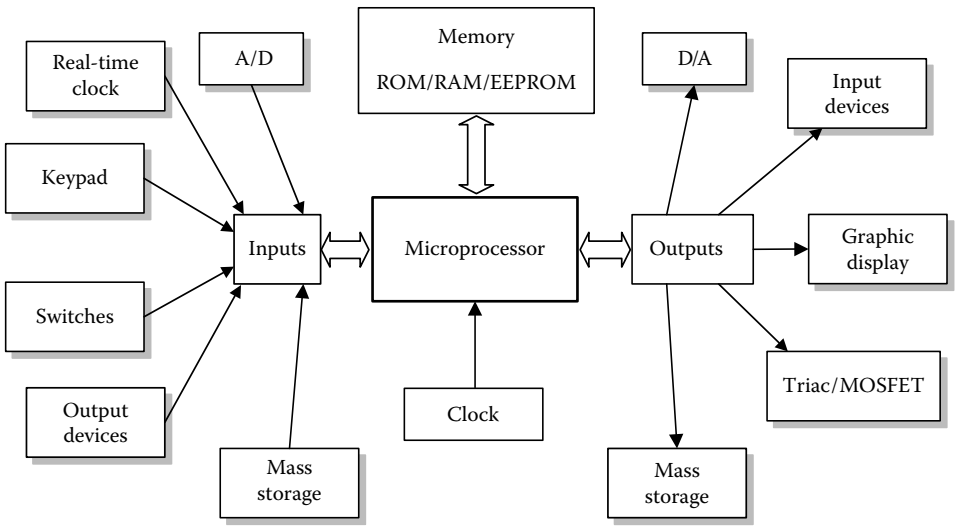


FIGURE 9.67 Conceptual diagram of a microprocessor-based system.

voltage in the range 0–1.5. Voltages in the range 1.5–3.5 will produce indeterminate results. Digital hardware using other voltage levels is also available. Of particular interest for portable applications are microprocessors using 3 V (or even 1.5 V) and 0 V. Unless explicitly noted, all signals are voltages as they vary with respect to a common reference.

Taken together, n-bits that are labeled, for example, with $d_{n-1}, d_{n-2}, \ldots, d_1, d_0$ can have any one of 2^n different binary combinations or binary assignments. A group of $n = 8$ bits is called a byte, which can take on any one of 256 different binary assignments. Another commonly used quantity is $2^{10} = 1024$, which is denoted by 1K. Thus, for example, we get $16\text{K} = 2^{14} = 16{,}384$.

Inputs can be analog in origin, and the microprocessor receives the result of sampling and analog-to-digital (A/D) conversion of an analog signal. Usually, analog signals are electrical signals generated by transducers that convert physical quantities of interest to voltages, such as, for example, speed of a motor, weight of an object, sound of someone speaking, stress of material under test, temperature in a building, electrocardiogram of a patient, acceleration of some structure, wind direction, oil pressure in an engine, etc. The analog signal can be the output of a receiver in a wireless communication application. The possibilities are endless. Thus, a microprocessor (the program it is executing) can receive information about time and the state of the world in which it is intended to operate.

The memory module of Figure 9.67 serves several purposes. First of all, it contains the binary instruction codes of the program that is executed by the microprocessor. The program instruction codes are another kind of input to the microprocessor. Memory may also be used for temporary storage and later retrieval of data produced by the program as it executes the algorithms that process the inputs. And, memory may contain various kinds of data tables that are referenced according to the instructions of the executing program. In contrast to the memory module, which must respond in time comparable to microprocessor speed of operation, slower mass storage devices also provide inputs that may be programs that will be transferred to memory for subsequent execution or that may be data that will be utilized by an executing program.

An important feature of using a microprocessor is that the functionality of an application is achieved mainly with software and the interaction through hardware between the software and the application environment. From a mass production viewpoint, this is almost a matter of trading hardware, a repeated expense, for software, a one-time expense, whenever possible. Thus, we are not concerned to design a general-purpose computer, but instead, we want to design the hardware that is sufficient to support the software that comprises the functionality of an application. Furthermore, using a microprocessor can achieve more than this, because with programmed hardware we can realize fundamental tasks such as conditional checking, storage, waiting, arithmetic, and timing, to mention a few, and higher level tasks such as complex algorithm calculations, decision making, control of machines, management of data, complex graphical interaction with the system user, and so on. When programmed properly, a microprocessor provides the ability to perform tasks that are difficult or impossible to achieve with any other kind of means or hardware.

The outputs from the microprocessor in Figure 9.67 are all digital signals. Like inputs, the meaning and destination of outputs can also be very diverse. Outputs can be one bit quantities that are intended to turn on and off, for example, light and sound indicators and devices such as, for example, TRIACs that can control AC power supply to equipment like AC motors, lights, heaters, and transformers, and, for example, power MOSFETs that can control DC power supply to equipment like solenoids, relays, DC motors, and lights. Outputs can be complex data structures that support graphical interaction between the user and the microprocessor-based system. As for inputs, here too the possibilities are endless. Outputs may be necessary to control the input devices or inform them of the readiness to receive inputs.

Outputs can also be 8, 16, or other multi-bit quantities that are, for example, printer control and text character codes, data to be placed in mass storage, a binary value to be transferred to a digital-to-analog (D/A) converter to synthesize an analog signal, data to be placed in memory, codes of characters to be displayed, data to be wirelessly communicated, protocols to establish communication links, and many other possibilities.

The clock module produces a voltage pulse train, called the system clock. Since microprocessor activity occurs due to level changes of the system clock, the speed at which the microprocessor executes instructions depends on the frequency of this pulse train. For example, if some particular microprocessor can fetch and execute an instruction in 5 clock cycles, and it can be clocked with a frequency of 250 MHz, then the microprocessor can execute the instruction in 20 ns. For most microprocessors, the number of clock cycles required to fetch and execute an instruction varies with the kind of instruction.

With all of the above modules interconnected to achieve a particular task, it will be the activity of the microprocessor, which is determined by the software that it executes, that controls the acquisition of inputs, processes the inputs, implements the intended functionality of the system, and generates and controls the placement of outputs.

Figure 9.68 shows a more functional structure than Figure 9.67 of a microprocessor-based system. An important feature of this structure is the use of a set of common communication paths, called the system bus, between all major modules. To identify any particular device, the microprocessor places a binary number, called an address, on the m-bit address bus. Actually, the electronic circuitry that comprises the microprocessor sets the voltages on the m address bus conductors to some combination of 0 and 5 V. There are 2^m different addresses that can be placed on the address bus. Every other module receives the address, and only the particular device that recognizes its own address, which is assigned and incorporated by the system designer, is supposed to respond. The data bus simultaneously transfers n-bits of data from the microprocessor to any device or from any device to the microprocessor. There are microprocessors with the number of address bits ranging from $m = 10$ to $m = 64$ and the number of data bits ranging from $n = 4$ to $n = 64$. We will work with microprocessors that produce an $m = 16$ bits wide address, which permits $2^{16} = 65{,}536 = 64\text{K}$ different addresses, and that use an $n = 8$ bits wide data bus. The control bus informs all modules about the kind of activity that is presently occurring or is about to take place on the system bus.

Usually, the different modules in the system have different time and electrical and operating characteristics. It is the purpose of the interfaces to make compatible the modalities of these different modules. Often, it is interface design that is the focus of an overall hardware design effort. Depending on the kind of control signals provided by a microprocessor and the kind of control signals that are preferred or required on the control bus, a microprocessor may also require an interface circuit to the system bus.

There are numerous different standard system bus definitions. For example, the public availability of the definition of the IBM PC bus has permitted the connection of innumerable different accessories, produced by many different manufacturers, to this bus. Such accessories and associated software were designed to be compatible with the definition of this bus standard. Over the past three decades, such standardization has contributed to the wide success of the PC in the home and in the workplace.

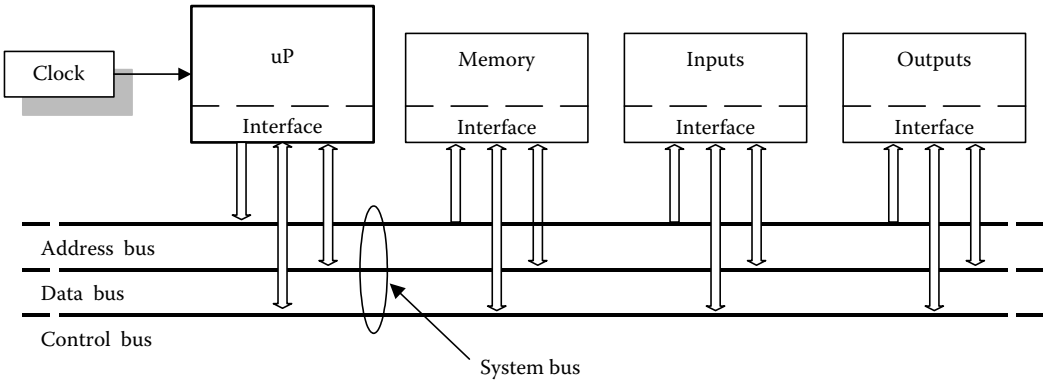


FIGURE 9.68 System block diagram.

However, in a dedicated application, adherence to a bus standard may not be necessary, particularly when there is no need to interconnect standardized modules. Here, we use the particular control signals of the microprocessor to form the system control bus as needed.

9.4.3 Memory

Almost all of the internal and external activity of a microprocessor depends on the ability to store, transfer, and retrieve data. A designer must understand how these operations take place. By way of presenting some memory devices, some terminology and supportive hardware are also introduced.

A circuit for a memory cell is shown in Figure 9.69. It is also called a binary cell (BC). Here, the activity of the flip-flop is determined by the logic signals at the points labeled *S*, for select, and R/$\overline{\text{W}}$, for read or write. The small circles attached to the OR gates and the buffers are called bubbles, and they mean logical complement. Thus, the buffer followed by a bubble is a NOT gate, and the flip-flop is made with two NOR gates. Another commonly used notation for logical complement is a small right triangle, as shown in the figure.

Note that we are using positive logic so that logic 1 means that a signal is high (5 V) and logic 0 means that a signal is low (0 V). When an operation or activity is enabled by taking a control signal high(low), then we say the control is active high(low). More briefly and without regard to the required level, an operation is enabled by asserting (or activating) its control signal(s).

The read and select control inputs of the BC are active high, while the write control input is active low, which is indicated by the bar over the *W*. To store a bit in the BC, the sequence of operations is (1) set the input, (2) set the R/$\overline{\text{W}}$ control signal low, and (3) momentarily select the cell. And, the write operation is complete.

A set of BCs that can be read or written in parallel is called a register, and the set of bits that a register can hold is called a word. A four words by 3 bits/word memory module is shown in Figure 9.70. The $m = 2$ bits input a_1a_0 is decoded with the 2×4 decoder, which has its own active low enable control. When enabled, the decoder output selects, according to its inputs, just one register (row of BCs) for a read or write operation. The binary assignment to a_1a_0 is called the address of the word to be referenced. For a read operation, the OR gates receive the bits of the selected word, while all other OR gate inputs are logic 0. For a write operation all BCs in a column receive the same input bit, while only the BC in the selected row accepts it. With additional columns, we can increase the register length to *n*-bits/word, and with an $m \times 2^m$ decoder we can have a memory size of 2^m registers, while only requiring *m* address bits to specify any particular memory register that is the object of a read or write operation (memory reference).

Figure 9.70 implies that the system data bus must consist of an *n*-bit unidirectional input bus to transfer data from memory to the microprocessor and an *n*-bit unidirectional output bus to transfer data from the microprocessor to memory. Practically, to reduce the number of communication paths, an *n*-bit bidirectional system data bus is preferred. This can be accommodated by connecting to each memory I/O bit pair, say the *i*th pair, an arrangement of tristate buffers as shown in Figure 9.71, where d_i

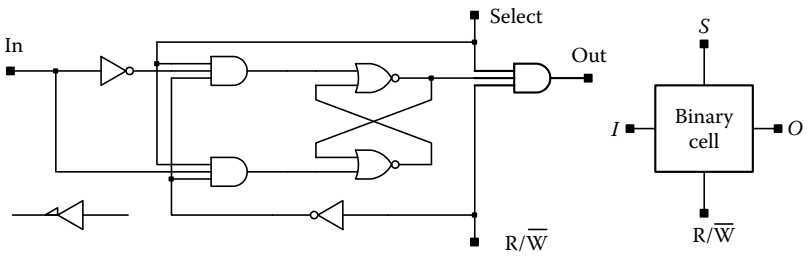


FIGURE 9.69 Circuit of a BC with select and R/$\overline{\text{W}}$ control.

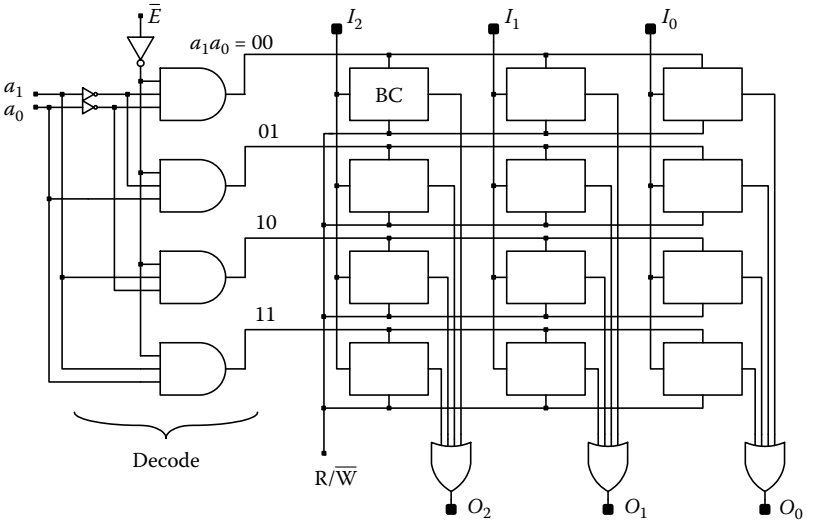


FIGURE 9.70 A multiword read/write memory circuit.

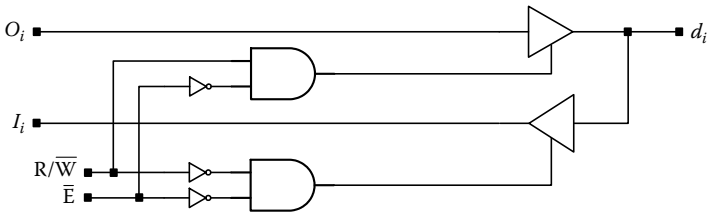


FIGURE 9.71 Unidirectional to bidirectional bus conversion.

is connected to the *i*th bit of the system data bus. When both buffers are in tristate (high impedance state), which occurs when $\overline{E}$ is high, then d_i is independent of O_i and I_i. Thus, the memory module can be electrically disconnected from the system bidirectional data bus. When $\overline{E}$ is asserted, then R/$\overline{W}$ controls the direction of data transfer. Figure 9.72 shows a more concise notation for the entire memory module. Here, the two address lines are grouped into one bus, and the three bidirectional data lines are grouped into another bus.

Since there is no particular required address sequence for reading or writing data to this memory, it is called random access memory (RAM). Since stored data remain intact as long as power is supplied to the circuit, this memory is called SRAM, contrary to another memory type, called dynamic RAM (DRAM),

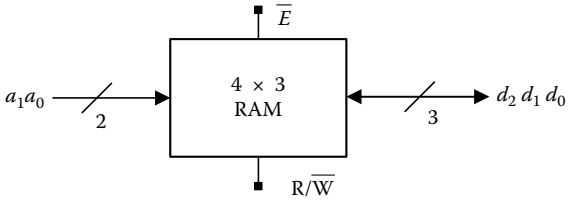


FIGURE 9.72 A memory module.

where the logic content of a BC is not based on the state of a bistable circuit, but is instead based on the presence or absence of charge on a cell capacitor. Since charge on capacitors dissipates, DRAM must be periodically refreshed, which requires additional control circuitry. However, DRAM can be fabricated to achieve much higher bit densities than of SRAM, resulting in the substantially lower cost of DRAM.

RAMs in various dimensions and package types are available. Commonly available SRAM chips range in size from 256×1 to $1024\text{K} \times 8$ (a 1 megabyte RAM), and the function and number of the control signals also vary a little. Commonly available DRAM chips range in size from $16\text{K} \times 1$ to $512\text{M} \times 1$ (a 512 megabit RAM). Important parameters for each memory type are the read and write cycle times, which must be less than the width of read and write time windows allowed by the devices that will reference this memory.

Larger RAM modules are constructed by the cascade and parallel connection of smaller RAMs. This is illustrated in Figure 9.73, which shows a $32\text{K} \times 8$ RAM made from $16\text{K} \times 4$ RAMs. If $\overline{E}$ is asserted, which takes the entire module out of tristate, then address bit a_{14} enables either RAMs 1 and 2 or RAMs 3 and 4 to access the data bus. The remaining address bits select a particular 4 bit register within each $16\text{K} \times 4$ RAM. Here, RAMs 1 and 3 hold the upper half of each data byte, while RAMs 2 and 4 hold the lower half of each data byte.

Whenever power is removed from the circuits of SRAM or DRAM, the data content is lost, and such memory devices are said to be volatile. To provide software for execution immediately after power is applied to a microprocessor, ROM is used. Figure 9.74 shows eight words by 2 bits/word ROM that has been, for example, programmed with the data given in the table. By opening, which is indicated with an *x*, the connections called links the ROM is programmed. These links remain open through power-down and power-up cycles, and thus, a ROM is said to be nonvolatile. The tristate buffers are controlled by the OE signal so that the ROM output can be electrically disconnected from a system data bus, which may then be used for transfer of data among other devices. An important parameter of a ROM is its read cycle time.

By providing a manufacturer a table of contents, a ROM can be programmed at the time of fabrication. Blank (unprogrammed) ROMs are available that can be one time programmed (OTP) in the field. These are called PROMs for PROM. There are also PROMs that can be erased by exposing the IC to a certain level of ultraviolet light for several minutes. Erasing is made possible by using an IC package with a clear

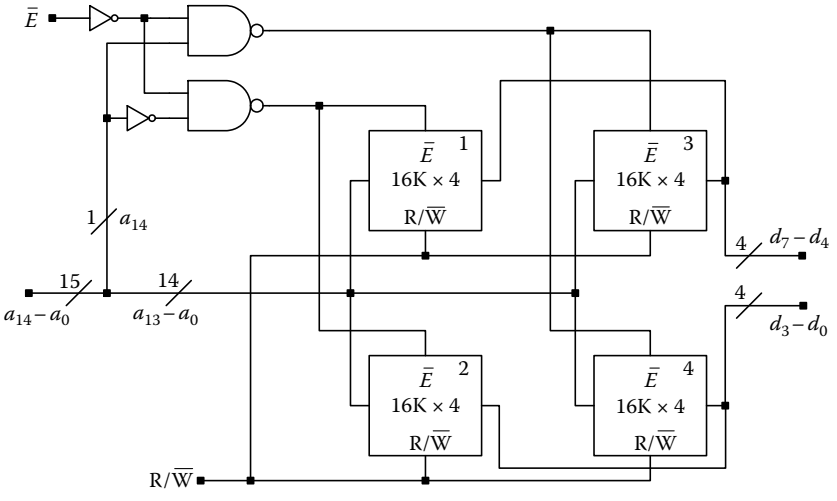


FIGURE 9.73 Cascade and parallel construction of a RAM module.

Address			Data	
a_2	a_1	a_0	O_1	O_0
0	0	0	0	1
0	0	1	1	1
0	1	0	0	0
0	1	1	0	1
1	0	0	1	0
1	0	1	1	1
1	1	0	1	1
1	1	1	0	0

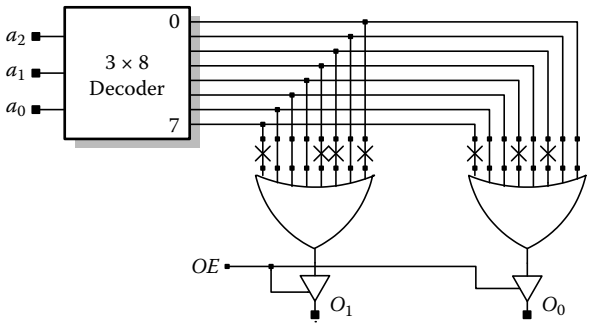


FIGURE 9.74 An eight words by 2 bits/word programmable ROM.

window just above the IC. These are called EPROMs for erasable PROM. After an EPROM has been erased, it can be programmed again.

Another EPROM type is the EEPROM that can be electrically erased by essentially overwriting previously stored data. However, the write cycle time is significantly longer than the write cycle time of conventional RAM. Commonly available PROMs and EPROMs range in size from 2K × 8 to 1024K × 8 (a 1 megabyte ROM). Commonly available EEPROMs are not as large as other ROMs, and they cost more than other ROM types. Like RAM modules, larger ROM modules can be made by combining smaller ROMs in a manner illustrated in Figure 9.73.

A recently developed memory type is the flash EEPROM, or flash memory. Its fabrication methodology is relatively inexpensive, and yields high-density memory modules having large data storage capacities. To rewrite to it, it must first be erased in bulk. These memories are used in, for example, digital cameras and portable flash mass storage devices.

Figure 9.75 shows an 8K × 8 EPROM and the package pin assignment. To place a byte into a particular ROM register, or program the ROM, the register address and intended content must first be supplied at

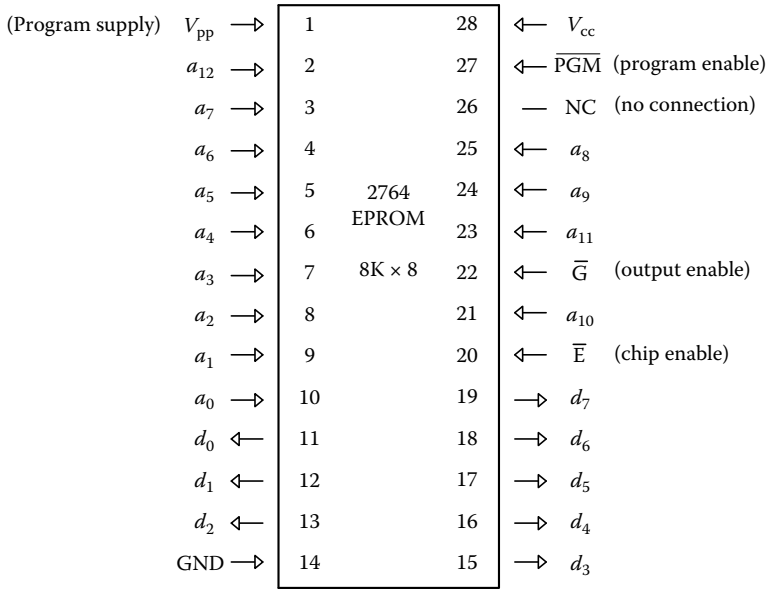


FIGURE 9.75 Pin assignment of a typical (the 2764) EPROM.

the address and data pins, respectively. Then, in addition to V_{CC}, the V_{PP} supply must be provided, the chip enable is asserted, and the $\overline{PGM}$ input is activated for a particular time duration. An instrument called a PROM programmer is used to do this. Some PROM programmers work by connection to a computer, and, in conjunction with software running on the computer, it can program a great variety of PROMs, which are selected from a menu of those PROMs supported by the software, with code in a file that it receives from the computer.

A programmed ROM can serve several purposes. Most computers have a ROM that contains program code that is always executed at power-up. Typically, this software, or firmware, is a utility (called a boot) that transfers software from a mass storage device to RAM. Then, after this transfer is complete, or perhaps after some additional ROM resident initializing code has executed, the software in RAM starts to execute. In a dedicated application, such as, for example, the microprocessor-based systems that control a car engine, it is more likely that ROM will contain all of the application software ready for use whenever the system is powered-up.

ROM is also used for other purposes. For example, by connecting three variables to the address inputs of the ROM given in Figure 9.74, this ROM can, through programming, be used to realize any two Boolean functions of three variables. ROMs are used to hold tables of data. For example, multiplication can be performed through table look-up. If the 4 bit binary codes of two digits that we want to multiply are used to form an 8 bit address of, for example, a 256×8 ROM, then the upper and lower halves of the retrieved data could each be the 4 bit binary codes of the product digits. The idea is that the address is formed with the input(s) of an operation (or argument of a function), and the precomputed and stored data are the output of the operation (or value of the function). EEPROMs are often used to hold tables of data that are generated during program execution, which must be available after a power-down and power-up cycle. A battery-backed RAM can also be used for such purposes.

Another way to use a ROM is for signal generation. By using a counter to supply an address sequence to a ROM, each bit of the data as it is clocked out of the ROM could be used as a control signal of some process. The frequency of the clock that drives the counter determines real time, and the variation between 0 and 1 of any particular data bit determines the resulting control signal shape. On the other hand, if the data word sequence coming from a ROM is inputted to a D/A converter, then a ROM can be used to generate an arbitrarily shaped analog signal that is repeated each time the counter repeats its count sequence.

The kind of memory that is used, its size, and the addresses to which any particular memory will be responsive must meet the requirements of an application. This information is often given in the form of a system memory map. The size of this map is the range of addresses that a microprocessor can specify.

With $m = 16$, an address in binary is denoted by $a_{15}a_{14}\ldots a_1a_0$. The most significant address bit a_{15} splits the entire 64K word memory space into two 32K word blocks, and 0xxxxxxxxxxxxxxx, where x means do not care, is any address in the lower 32K word block, while 1xxxxxxxxxxxxxxx is any address in the upper 32K word block. Similarly, address bits $a_{15}a_{14}$ together split the entire memory space into four 16K word blocks. If $a_0 = 0$, then the address is an even number, while if $a_0 = 1$, then it is an odd number.

Sometimes, it is more convenient to use hexadecimal notation. The symbol \$ will be placed in front of all numbers written with hexadecimal notation. Thus, \$XXXX is any 16 bit number in hexadecimal notation, while \$XX is any 8 bit number. As the most significant hexadecimal address digit changes by \$1, the address changes by 4K.

Suppose a design must include 8K bytes of EPROM starting at address \$0000, 1K bytes of EEPROM starting at address \$4000, and 32K bytes of RAM starting at address \$8000. The ROM, EEPROM, and RAM must be responsive to addresses in the ranges \$0000–\$1FFF, \$4000–\$43FF, and \$8000–\$FFFF, respectively.

Since the three most significant address bits split the memory space into eight 8K word blocks, the ROM should be enabled when these address bits are $a_{15}a_{14}a_{13} = 000$, and then the remaining address bits $a_{12}a_{11}\ldots a_1a_0$ covering an 8K word space are the address input to the ROM. Similarly, the EEPROM

should be enabled when $a_{15} \ldots a_{10} = 010000$, and the remaining address bits $a_9 \ldots a_1 a_0$ covering a 1K word space are the address input to the EEPROM. The RAM should be enabled when $a_{15} = 1$.

Thus, the size and desired position of a memory device determine how to place it in the memory space. The other important aspect of interfacing memory to the system bus and eventually to a microprocessor is that a memory module must be responsive to control signals issued by the microprocessor to present and accept data in certain particular time windows. This will be considered further when we look at the timing of microprocessor control signals. An alternative mode of operation would be for the microprocessor to automatically wait for a response signal from memory whenever it references memory before completing the memory reference cycle.

A memory map for this design is shown in Figure 9.76. Here we see the size, kind, and position in the memory space of the microprocessor of actual memory devices. When possible, it is also useful to describe the location and purpose of particular program code modules.

Assuming that the EPROM contains program code, some example instructions have been placed at the beginning of this memory. These are binary numbers that are machine instruction codes for the Z80 microprocessor. The Z80 microprocessor is a type of processor that after being reset, fetches instruction codes starting at address $0000. The first instruction, i.e., $ED $56, is the Z80 machine code that selects its method 1 for responding to interrupts. Another commonly used notation for hexadecimal is to attach the suffix H. The next instruction, i.e., 31H, is the machine code that loads the stack pointer (SP) register with an address given by the next 2 bytes, i.e., $F000. These instructions could be the part of a program that initializes the processor. To understand how an instruction is processed, it is useful to study the architecture of a microprocessor.

For programming convenience, programs are usually written using mnemonics of instruction codes that are indicative of instruction activity. The set of mnemonics and associated notational convention for all of the instruction codes that a particular microprocessor can execute form the assembly language of the microprocessor. Different microprocessors, with different instruction sets have different assembly languages. Furthermore, different manufacturers adopt different mnemonics

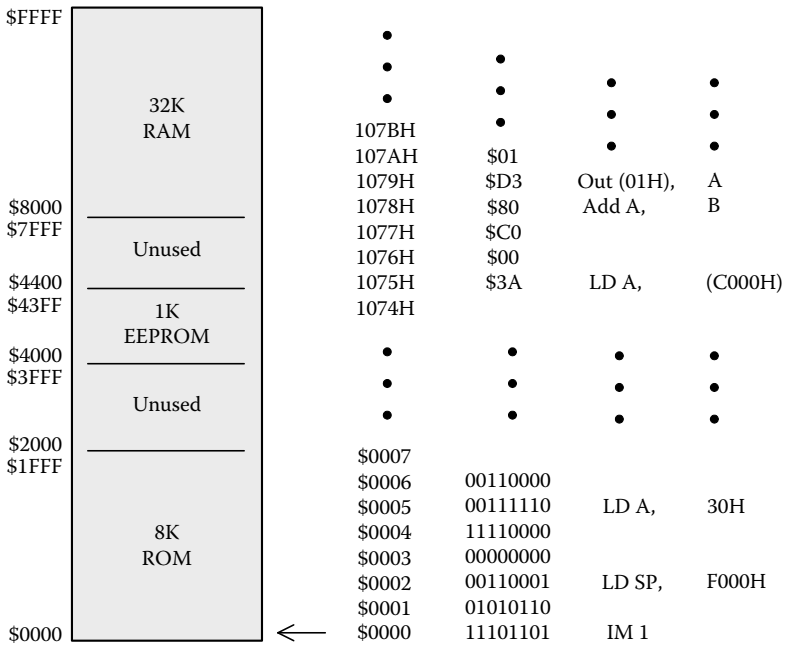


FIGURE 9.76 Memory map, sample ROM content, and assembly language source.

for their microprocessor instruction sets, causing another variation among assembly languages. However, in principle, there remain many common attributes from one assembly language to another. A program written in assembly language must be converted into machine code for storage and eventual execution. A program that performs this conversion is called an assembler. From the viewpoint of the assembler, an assembly language program is an input character string, and the machine code output is another character string.

9.4.4 Microprocessor Architecture

Employing a microprocessor in a dedicated application does not require detailed knowledge about its internal behavior. However, it is useful to have some insight about how instructions, i.e., their codes, are processed (executed) by the hardware. Understanding the relationship between software and hardware can affect the selection of a particular microprocessor, and the design process. Moreover, the hardware designer should also understand the programming model of a microprocessor.

From a programmer's viewpoint, a microprocessor is defined by its programming model and its instruction set, which together comprise the architecture of the microprocessor. The programming model consists of the set of internal registers that are involved in the execution of operations as specified by the instruction set. These registers do different specialized tasks. The purpose, capability, and number of these registers can vary greatly from one microprocessor to another.

Basically, within the programming model, microprocessors have four kinds of registers. There are address registers that are used to form and hold addresses to be used for referencing memory and other devices to obtain program instruction codes and their operands and to specify source and destination memory locations for data read and write operations. There are data registers that can be the source of data for an operation or the destination for the result of an operation. There are operational registers that have associated hardware to perform, for example, logical and arithmetic operations. And, there are status/control registers that configure the operation of the microprocessor and support different kinds of conditional instructions.

An operational register possessed by most microprocessors is the accumulator. It is commonly denoted by register A. Figure 9.77 illustrates how an A register functions within a microprocessor. Associated with reg. A are the ALU and the condition code register (CCR) or status register (SR). Inputs to the ALU can come from reg. A, reg. B, and the CCR, and its activity is determined by the function select word $f_{k-1} \ldots f_1 f_0$. Reg. A receives its input from the ALU, and reg. B receives its input from the n-bit internal data bus. The activity of this circuit is determined by the control signals that are applied at all of the points labeled with triangles. For example, by asserting the E, for enable, input of reg. A, the n-bit word coming from the ALU is latched (loaded) into reg. A. The content of reg. A can be placed on the internal data bus by asserting the control signal of the n tristate buffers connected to its output. The content of the internal n-bit data bus can be latched into reg. B by asserting its E control input.

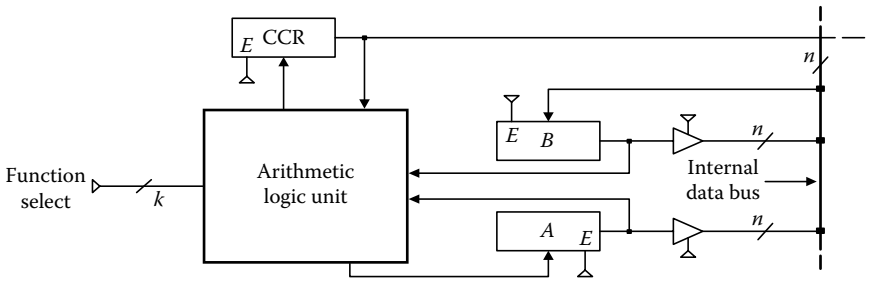


FIGURE 9.77 Register-to-register transfer activity of an accumulator.

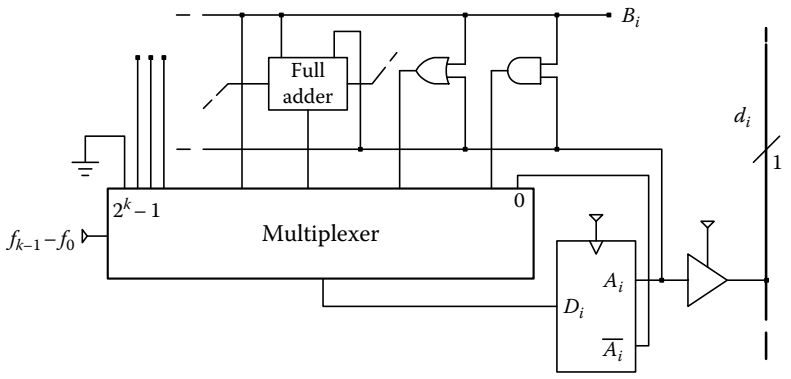


FIGURE 9.78 Bit level activity of an ALU.

Figure 9.78 illustrates how the ALU performs its task at the gate (bit) level. Notice how the function select lines determine, as they do for the MUXs of all the other bits, which MUX input is latched into the *i*th flip-flop of the accumulator when its *E* control input is asserted. For example, if $f_{k-1} \ldots f_0 = 0 \ldots 0$, then the accumulator will be complemented, as if the microprocessor has just read in the machine code for the complement accumulator instruction, which has the Z80 assembler mnemonic CPL. Thus, the machine codes of instructions such as ADD A, B; RLA; INC A; LD A, 80 H; etc., eventually determine the binary assignments of these function select lines to accomplish the instruction tasks. There can be as many as 2^k different operations involving the accumulator that can be achieved by this ALU structure.

All of the control bits required in Figures 9.77 and 9.78 come from memory called control ROM or microstore. Figure 9.79 illustrates the data paths of a microprogrammable microprocessor. Each word, consisting of perhaps 16–128 bits, in control ROM is called a microinstruction. The busing of microinstructions throughout the microprocessor is not shown in Figure 9.79. Instead, these interconnections are indicated by labeling with the small triangles.

Figure 9.79 also illustrates a level of design called the register transfer level. There is still another level of greater detail before we get to a level of IC design detail that shows the interconnection of transistors, resistors, diodes, and conductors. This is the logic gate level. However, at that level, the detail of design would probably be too much and detract from an understanding of microprocessor operation. Here the intent is to exemplify how hardware processes an instruction code.

Each microinstruction is partitioned into a set of fields, and one of these fields holds the accumulator function select line assignment. Other fields hold data that signify: (1) which registers are the source of data for the internal and external buses; (2) which registers are the destination of data on the internal and external buses; (3) register activity control such as clear, increment, decrement, load, and tristate; (4) a *j*-bit address of the next microinstruction; and (5) the external control word that informs external devices of the present microprocessor activity. Each microinstruction coming out of control ROM can cause many activities to take place at the same time within the microprocessor. The other registers in the diagram of Figure 9.79 perform the following tasks.

MAR—Memory Address Register. It holds the address that the microprocessor can place on the external address bus, and it is loaded from the internal address bus by asserting its enable input.

PC—Program Counter. This register holds the address of the next instruction or instruction operand. It can be cleared, loaded, and incremented by asserting appropriate enable inputs. To fetch program code (either an instruction code or instruction operand), its content is transferred to the MAR. Usually, after each time its content has been used to fetch a byte of program code, its content is incremented.

SP—Stack Pointer. It holds an address that points to RAM that can be used for temporary storage. It can be loaded, incremented, and decremented by asserting appropriate enable inputs. Certain

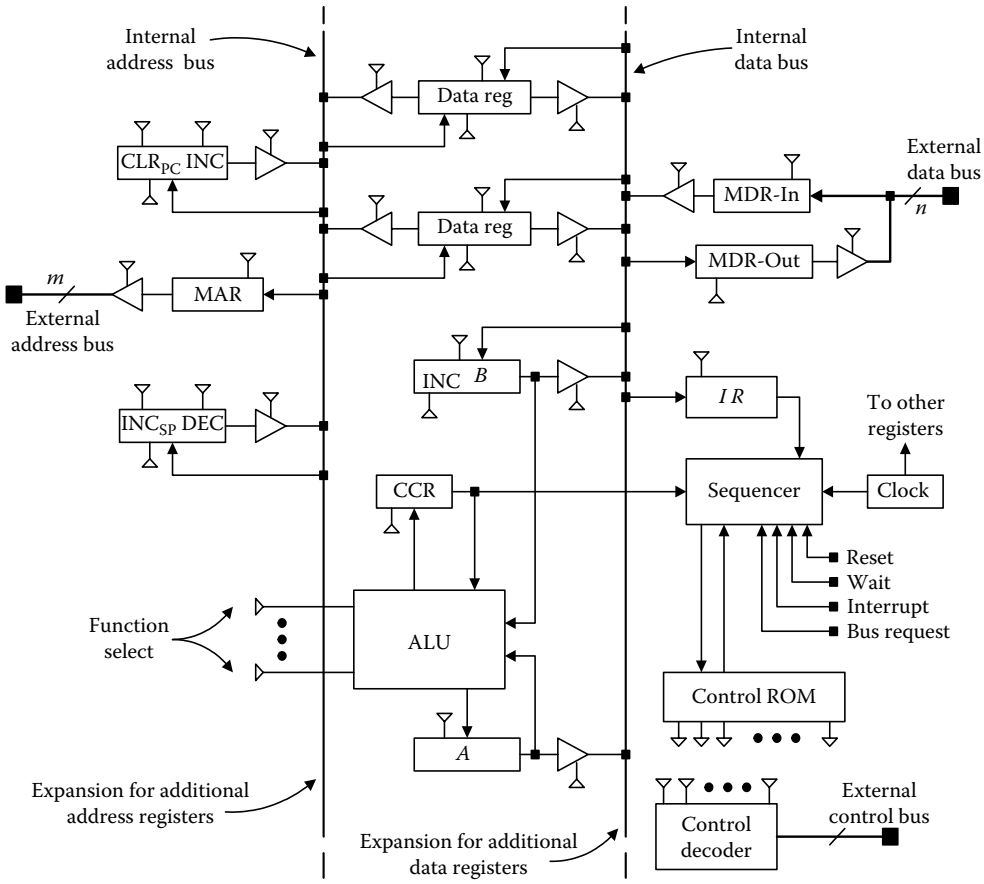


FIGURE 9.79 Data paths of a microprogrammable microprocessor.

instructions can cause its content to be transferred to the MAR for a memory read or write operation. The RAM that is referenced with the SP is called the stack. Typically, the SP is implicitly decremented before (after) writing to the stack, and it is implicitly incremented after (before) reading from the stack. Thereby, the stack is a last in and first out memory area that is used to support, among other things, subroutine calls.

MDR-IN—Memory Data Register IN. It receives its input from the external data bus, and it can drive the internal data bus.

MDR-OUT—Memory Data Register OUT. It receives its input from the internal data bus, and it drives the external data bus.

IR—Instruction Register. When the microprocessor is executing an instruction fetch from external memory, this register receives the content of the MDR-IN register, which is then assumed to be an instruction code.

CCR.—Each bit in this SR is indicative of the result of some previous microprocessor operation. The meanings of the flags (bits) in the CCR vary from one microprocessor to another. Typically, this register contains: (1) a carry flag that is set or reset depending on whether or not the previous add or subtract instruction produced a carry or required a borrow out of the most significant bit of the arithmetic operation, (2) a zero flag that is set if the previous instruction produced a zero result and reset if the instruction result is not zero, (3) an interrupt enable flag that can be set or reset by an instruction to

allow for software control over whether or not the microprocessor can respond to maskable interrupts, and (4) other flags. Usually, a microprocessor's instruction set includes program flow control instructions that are conditioned on these status flags such that if the flag is set (reset), then the instruction is executed, and if the flag is reset (set), then the instruction is not executed (skipped).

Sequencer. The sequencer consists mainly of a ROM and some sequential logic. It uses the instruction code and SR flags to form an address to its own ROM from which is obtained a *j*-bit microinstruction address that is applied to the control ROM. Furthermore, in response to asynchronous external control signals, such as reset, interrupt, bus request, wait, and others, it generates addresses of microinstructions that cause activity appropriate for these inputs.

Control Decoder. Depending on the input coming from control ROM, it generates external control signals that are intended to be used to inform external devices about the present activities of the microprocessor and to synchronize activities on the external address and data buses.

Data Register. This register is used for temporary storage of data. Usually, there are several registers like it. Some can also drive the lower byte of the internal address bus, while others can also drive the upper byte of the internal address bus. Still others can receive data from the upper or lower byte of the internal address bus. Some may have low-level arithmetic/logic capabilities.

The execution of each microprocessor instruction involves the execution of a set of microinstructions, called a microprogram. Thus, the functionality of a microprocessor's instruction set is determined by the microprograms stored in control ROM. After system reset or the completion of each instruction, unless an external asynchronous input is active, the microprogram that performs an instruction fetch is executed. It uses the content of the PC to point to the instruction, and it increments the PC so that the PC points to an instruction operand or the next instruction. Then, depending on the instruction code and SR content presented to the sequencer, a particular microprogram is executed to complete execution of an instruction. If an instruction code requires that operands be fetched, then the PC is further incremented so that, after an instruction has executed, the PC is pointing to the next instruction.

Microprograms for program flow control instructions such as JUMP or BRANCH elsewhere cause the PC to be loaded with the address operand of the instruction. Moreover, microprograms for instructions such as CALL or BRANCH subroutine first cause the SP to be used for storing in the stack the content of the PC before loading the PC with the address operand. Then, the microprogram for an instruction such as RETURN, which is used to terminate a subroutine, causes the SP to be used for retrieving from the stack the address of the instruction following the CALL subroutine instruction, which is then loaded into the PC.

This architecture can be expanded to include additional address registers and other special purpose address registers such as index registers, another PC and SP, additional accumulators and data registers, and so on. And, the widths of the address and data bus can be increased, contingent upon fabrication issues. Moreover, algorithmic instruction types can be supported since executing the desired activity of an instruction is a matter of writing a microprogram to accomplish all of the required register-to-register transfer activities.

9.4.5 Design with a General Purpose Microprocessor

The Z80 is a general purpose microprocessor, and it is available in a 40-pin dual in-line package (DIP). It is an 8 bit machine, i.e., its data bus is 8 bits wide. It has also evolved into a great variety of similar microprocessors. From a software development viewpoint, this machine has numerous features, as an examination of its extensive instruction set and programming model, which is given in Figure 9.80, will show. It has an accumulator (A), status or flag (F) register, six data registers (B, C, D, E, H, and L) that can be paired for use as address registers (BC, DE, and HL), an SP, a PC, two index registers (IX and IY), a register (I) that is used by its indirect vectored interrupt processing method, and a 7-bit counter register (R) that can be utilized for DRAM refresh. Furthermore, it can quickly change programming context by swapping with the alternate (primed) register set. Among its over 600 instructions, there are data transfer, arithmetic, logical and rotate, branch (or jump), stacking, I/O, program control, exchange,

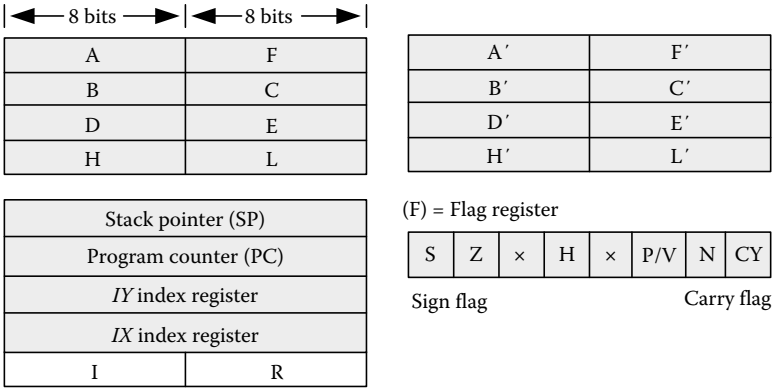


FIGURE 9.80 Z80 programming model.

block transfer, search, and bit manipulation instructions. Due to space limitations, the information given here is necessarily limited. Complete information can be found in textbooks or the data book from the manufacturer.

The pin assignment of the Z80 is shown in Figure 9.81. Power is supplied at pins 11 and 29, and a conventional crystal-controlled oscillator is used to provide the system clock ϕ at pin 6. If the load of each system address bus bit is only one or two TTL loads, then the outputs $a_{15}a_{14}\ldots a_1a_0$ can drive the system address bus. However, to accommodate a greater load, two unidirectional tristate octal buffers, as shown in Figure 9.81, can be used. These buffers also have hysteresis, which helps to produce sharper bus signals. Similarly, it is likely that the Z80 data signals $d_7d_6\ldots d_1d_0$ must be buffered to accommodate

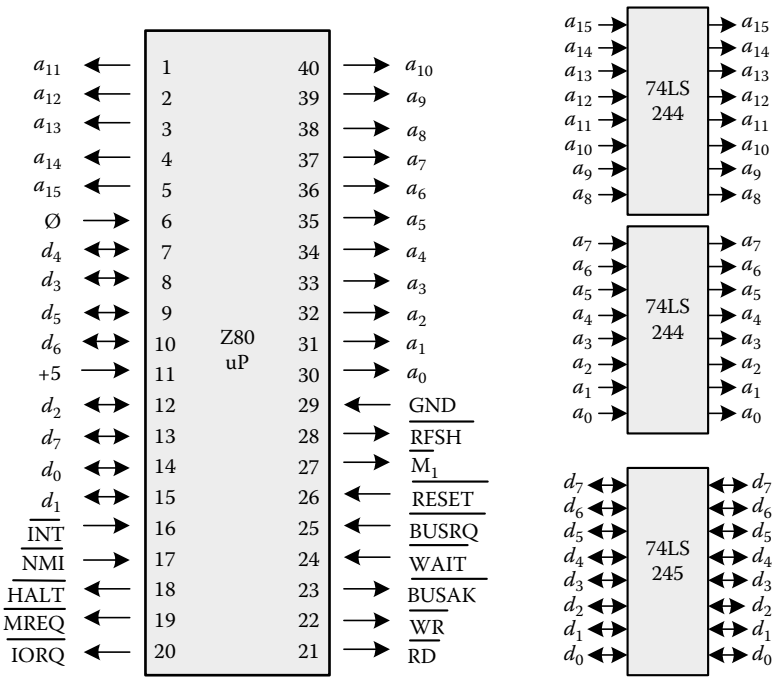


FIGURE 9.81 Pin assignment of the Z80 microprocessor.

system data bus loading. A bidirectional buffer, as shown in Figure 9.81, can be used for this. Z80 control signals must be used to control buffer direction. Also, define the system data bus by terminating it with pull-up resistors. The Z80 control signals perform the following tasks.

$\overline{\text{RESET}}$—This active low input resets the interrupt enable flag, clears registers I and R, causes all control signals to become inactive, sets the interrupt processing method to method 0, and clears the PC. When this input is released, the Z80 starts to fetch the first instruction code (op-code) from the memory register with address held by the PC.

$\overline{\text{MREQ}}$—This output becomes active whenever the Z80 is performing an operation, such as op-code fetch or instruction operand fetch that references an external device with a 16 bit address. It indicates that the address bus holds a valid address. The devices that respond are said to be positioned in the memory space.

$\overline{\text{IORQ}}$—This output becomes active whenever the Z80 is executing either an IN, for input, or an OUT, for output, instruction, both of which reference external devices with an 8 bit address that is placed on $a_7 \ldots a_0$. It also indicates that the address bus holds a valid 8 bit address. The devices that respond are said to be positioned in the I/O space, which is separate from the memory space.

$\overline{\text{RD}}$—This indicates that an external device must place valid data on the data bus, which the Z80 will soon accept.

$\overline{\text{WR}}$—This indicates that the Z80 is driving the data bus with valid data.

$\overline{\text{M}}_1$—This output is active while the Z80 is fetching an op-code. The only other time it becomes active is in response to a maskable interrupt.

$\overline{\text{HALT}}$—This output becomes active when the Z80 has stopped fetching additional instructions due to having executed a HALT instruction. The microprocessor can only continue instruction execution upon activation of an interrupt input.

$\overline{\text{WAIT}}$—This input, when active, causes the Z80 to hold constant its address and control signal outputs until this input is released (no longer active). When referenced, a slow memory or I/O device can cause the Z80 to wait by asserting the $\overline{\text{WAIT}}$ input until sufficient time has elapsed to allow the device to respond to the reference.

$\overline{\text{RFSH}}$—When active, this indicates that $a_6 \ldots a_0$ holds the content of counter register R, which, along with an active $\overline{\text{MREQ}}$ can be used to refresh dynamic memory.

$\overline{\text{NMI}}$—Whenever this nonmaskable interrupt input goes through an active low edge, an internal flip-flop is set, and other interrupts are disabled. The interrupt is said to be latched. At the completion of every instruction, this flip-flop is checked, and if it is set, then the Z80 first stacks the PC and then it loads the PC with \$0066, where the first op-code of an interrupt service routine must be located. This way of loading the PC in response to an interrupt is called a direct interrupt, and it is intended for an unconditional and fast response to, for example, battery low detected, temperature too high detected, or some other urgent situation, since it requires no further action from the interrupting device.

$\overline{\text{INT}}$—At the completion of every instruction the Z80 checks this maskable interrupt input. If it is active and interrupts have been enabled with the EI instruction, then interrupt processing commences according to the interrupt method stipulated by the IM i, $i = 0$, 1, or 2, instruction. If $i = 1$, for the direct method, the PC is first stacked, and then it is loaded with 0038H, the address used to obtain the first op-code of an interrupt service routine. If $i = 0$, for the vectored method, the Z80 acknowledges the interrupt by asserting the $\overline{\text{IORQ}}$ line while the $\overline{\text{M}}_1$ signal is active. This combined activity is then interpreted externally to be an interrupt acknowledge signal, denoted by $\overline{\text{INTA}}$. In response to the $\overline{\text{INTA}}$ signal, the interrupting device then has the opportunity to place the op-code for the one byte Z80 instruction RST, N on the data bus, where $N = 0, 1, \ldots, 7$. Then, the Z80 stacks the PC and loads it with an address given by 8 N. Thus, depending on N, a 3-bit code embedded within the RST instruction, the first op-code of

the interrupt service routine can be located in one of eight memory locations. If $i = 2$, for the indirect vectored method, the PC is first stacked, and then the Z80 acknowledges the interrupt by asserting the $\overline{\text{INTA}}$ signal. In response to the $\overline{\text{INTA}}$ signal, the interrupting device must then place a byte given by xxxxxxx0 on the data bus. The Z80 then uses this byte as the low address byte and the content of the I register as the high address byte to fetch two consecutive bytes from memory that are then loaded into the PC. Thus, the I register points to a 256 byte block of memory, where there can be 128 interrupt vectors one of which is selected by the byte provided by the interrupting device.

$\overline{\text{BUSRQ}}$ and $\overline{\text{BUSAK}}$—The first signal is an input that indicates to the Z80 that an external device wants to take control of the system bus. The Z80 completes execution of the present machine cycle, takes its address, data, and tristate control signals into tristate, and acknowledges the request with an active $\overline{\text{BUSAK}}$ signal.

Most of these signals are representative of the kinds of control signals that microprocessors have. Their usefulness becomes more apparent as we look at the timing of microprocessor activity and the impact this has on the design of hardware so that the microprocessor can accomplish read and write data transfers. Figure 9.82 shows the Z80 timing diagram for an op-code fetch, and Figure 9.83 shows timing diagrams for the other memory and I/O references. For the purpose of studying the timing of events, we do not have to know actual binary assignments on the address and data bus lines. Therefore, the

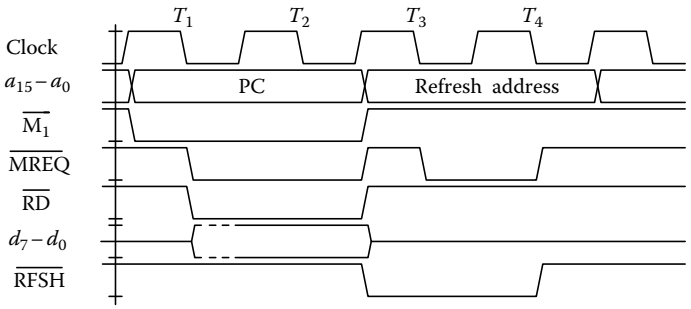


FIGURE 9.82 Z80 timing diagram for an op-code fetch.

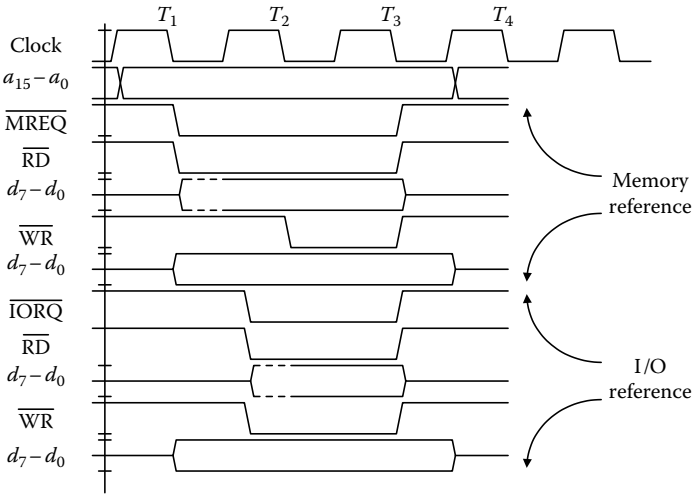


FIGURE 9.83 Z80 timing diagrams.

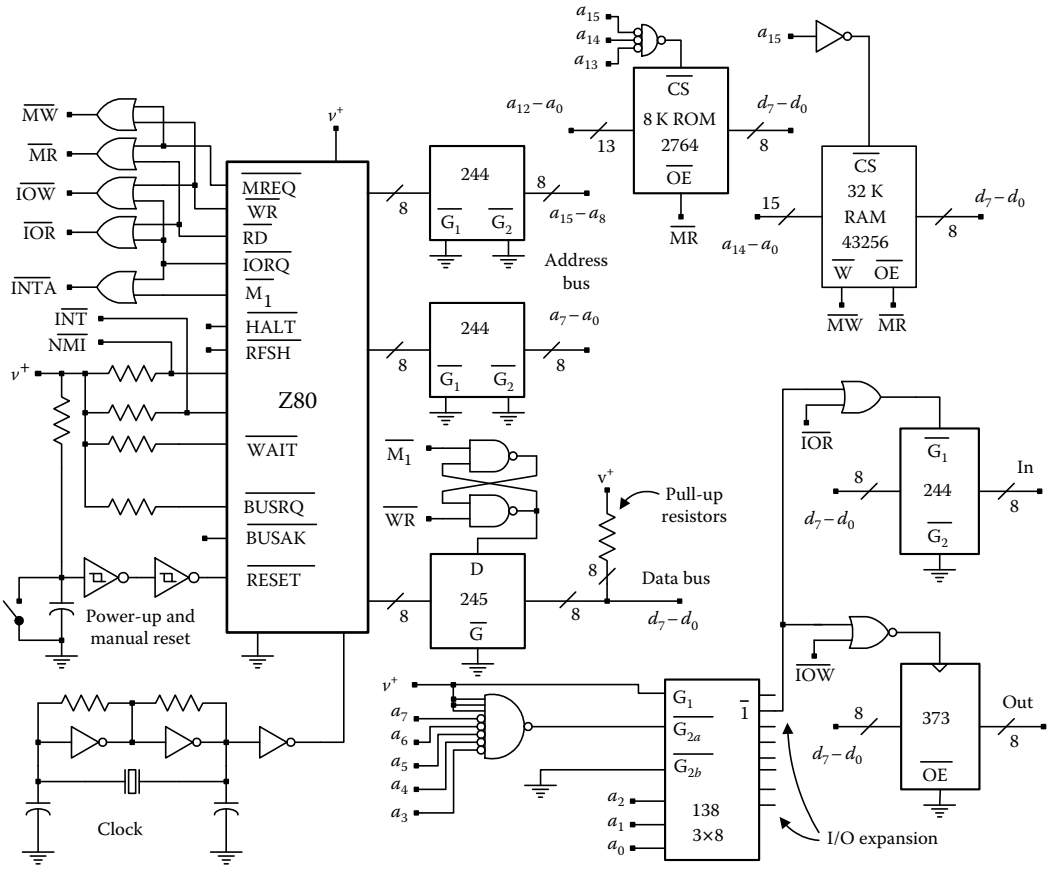


FIGURE 9.84 Z80-based microcomputer.

convention in these diagrams is intended to indicate when these signals either become relevant for the operation at hand or switch to the present binary assignment, whatever it may be, from some previous binary assignment.

As we look at the timing diagrams, it will also be useful to see how these signals are used. Figure 9.84 shows a schematic of a microcomputer designed according to the memory map of Figure 9.76. Whenever possible, labeling is used to indicate connections.

Depending on the instruction, the Z80 uses from 4 to 23 clock cycles to execute an instruction. A group of clock cycles within the execution time of an instruction that accomplishes a major activity is called a machine cycle. Each clock cycle within a machine cycle is labeled with T_i, $i = 1, 2, \ldots$, and all activities occur at either leading or trailing clock edges. To better understand such groupings, suppose that after executing many instructions that were retrieved from ROM, the Z80 is about to execute the next few instructions located in memory shown in Figure 9.76.

The PC is set to PC = 1075H, pointing to the op-code at memory location 1075H. This is the op-code for load accumulator with data using the two bytes that follow the op-code to form an address to point to the data, i.e., LD A (C000H). This is a 3 byte instruction, and, according to the Z80 manual, it requires 13 clock cycles, which are grouped into four machine cycles, to execute.

Referring to Figure 9.82, at the leading edge of the first clock cycle T_1 of the first machine cycle of this instruction the Z80 asserts the $\overline{M_1}$ signal, and it drives the external address bus with the address 1075H. At the trailing edge of clock cycle T_1, it asserts the $\overline{MREQ}$ signal and the $\overline{RD}$ signal. At the time that $\overline{RD}$

becomes active, the Z80's data bus input register is enabled to follow the content of the external data bus. At the leading edge of clock cycle T_3, the Z80's data bus input register is disabled from following the external data bus, and the content of this register is accepted as an op-code.

During the time interval from the trailing edge of T_1 until just before the leading edge of T_3, which is slightly less than 1.5 clock cycles, external hardware has the opportunity to place the addressed data, which will be processed as an instruction op-code, on the data bus. A designer must ensure that external hardware is fast enough to do this in a timely manner. If it is necessary to use an external device with an access time that is greater than this allotted time, then a counter, called a wait state generator, can be used. Once the slow external device detects (by address and control signal decode) that it is supposed to put valid data on the data bus, then it enables the wait state generator, which should be driven by the system clock, to activate the $\overline{\text{WAIT}}$ input of the microprocessor for a number of clock cycles that will give the device an opportunity to place data on the data bus. Each such clock cycle is called a wait state (labeled with T_w), and the number of wait states needed will depend on the access time of the device being referenced. If several external devices need to generate different numbers of wait states, then the outputs of the wait state generators of all such devices can be OR'd to present just one wait input to the microprocessor. Thereby, the machine can run as fast as each different external device will allow.

During the next 2 clock cycles, the received op-code is interpreted so that the Z80 becomes set to perform two more memory references, called machine cycles M_2 and M_3, to fetch each byte of the address operand. Machine cycle M_1 consists of 4 clock cycles, and while the Z80 is processing a retrieved op-code during clock cycles T_3 and T_4, the address bus is available for DRAM refresh. During machine cycle M_4, the address C000H, which was obtained during M_2 and M_3, is placed on the external address bus, and execution of this instruction is completed by transferring the content of memory location C000H to the accumulator. Machine cycles M_2 and M_3 each require 3 clock cycles, and machine cycle M_4 also requires 3 clock cycles. If whenever this RAM module is referenced, it issues wait states, then the actual number of clock cycles required to execute this instruction will be more than 13 clock cycles.

Notice that the op-code determines the kind and number of additional machine cycles that are necessary to complete execution of an instruction. It also determines how much the PC must be incremented. Thus, whenever execution of an instruction has been completed, the PC is pointing to an op-code of the next instruction. Furthermore, virtually all activity involves some kind of synchronized register-to-register transfer.

The op-code at location 1078H is the code for the 1 byte instruction, ADD A, B. Since this addition requires no additional memory reference, it requires 4 clock cycles (1 machine cycle) to execute. The next instruction, i.e., OUT (01H), A, is a 2 byte instruction, and it requires 3 machine cycles to execute. Since this is an OUT instruction, the Z80, during machine cycle M_3 does the following: (1) transfers the address 01 H, which it obtained during M_2, to the lower byte of the MAR to drive the lower byte of the system address bus, (2) transfers the content of the accumulator to the data bus out register at the trailing edge of T_1 to drive the external data bus, which, as shown in Figure 9.83, incurs a set-up delay, (3) activates the $\overline{\text{IORQ}}$ control signal at the leading edge of T_2, which incurs a set-up delay, and (4) activates the $\overline{\text{WR}}$ control signal, which also incurs a set-up delay, until the trailing edge of T_3. Thus, the I/O device that is supposed to receive the content of the accumulator has slightly less than 1.5 clock cycles during which it must capture the content of the system data bus. By decoding the address, which for an I/O device is called an I/O port address, and the control signals, an I/O interface can enable an external register to start to follow the data bus shortly after $\overline{\text{IORQ}}$ and $\overline{\text{WR}}$ have become active. Then, by the time that the $\overline{\text{WR}}$ signal becomes inactive, the external register should hold the content of the data bus.

9.4.6 Interfacing

The circuitry, or more broadly, the method that makes compatible the operations of devices so that these devices can exchange signals (data or codes) is loosely called an interface. The devices on opposite sides of an interface can be different in many ways. Interfacing two devices can encompass a variety of

requirements, such as, for example, (1) impedance matching, (2) voltage or current level conversion, (3) translation of control signal meanings, (4) protocol conversion, (5) signal timing alignment, (6) electrical isolation, (7) exchange of status information, (8) data format conversion, and (9) resolution of other incompatibility issues.

Figure 9.84 gives some examples of the goals of interface design. The unidirectional and bidirectional buffers make compatible the drive (current source) capability of the microprocessor and the drive requirements of the address and data bus. The particular control signals of the Z80 are interfaced to the system control bus with combinatorial logic to provide drive as well as more explicit control signals. The interfaces between the memory devices and the system bus utilize the address bus to position memory in the desired locations in the memory space. We must be certain that only one device can be a source to the data bus at a time. The interfaces also utilize control bus signals to produce the particular kinds of control signals required by the memory devices and to activate memory at times compatible with microprocessor timing.

Figure 9.84 also shows an input and an output port. The interface for this parallel I/O port also utilizes the address bus, and, instead of using control bus signals that go active due to memory reference instructions, it uses control bus signals that go active due to the IN and OUT instructions of the Z80. For output, it decodes the control bus and the lower half of the address bus to enable (clock) the octal latch to capture the content of the data bus at the right time and in response to the intended (its own) address $01. Thus, by executing an output instruction, like the one located at address $1079 in Figure 9.76, the octal latch receives the content of the accumulator. For input, the interface decodes the address and control bus to enable (take out of tristate) the octal buffer, which permits it to drive the data bus with its input at the right time and in response to the intended address. Thus, by executing an input instruction such as IN A, (01H), the accumulator will receive the byte at the buffer input. Depending on the address decode logic, the I/O port can be positioned anywhere in the I/O space of the microprocessor.

By using the $\overline{\text{MREQ}}$ control signal, instead of the $\overline{\text{IORQ}}$ control signal, and the entire address bus, I/O ports can also be positioned in the memory space of the microprocessor. The variety of conventional memory move instructions can then be used for I/O. Some microprocessors, like the M68HC11, do not have I/O space separate from memory space and instructions for I/O in addition to memory move instructions. Instead, no distinction is made between memory and I/O references. This eliminates the need for explicit control signals that interface hardware requires to distinguish memory and I/O references.

One or more parallel I/O ports can serve many different applications. With circuits like those given in Figure 9.85, a bit of an output port can turn on and off a light or sound indicator, or DC or AC power

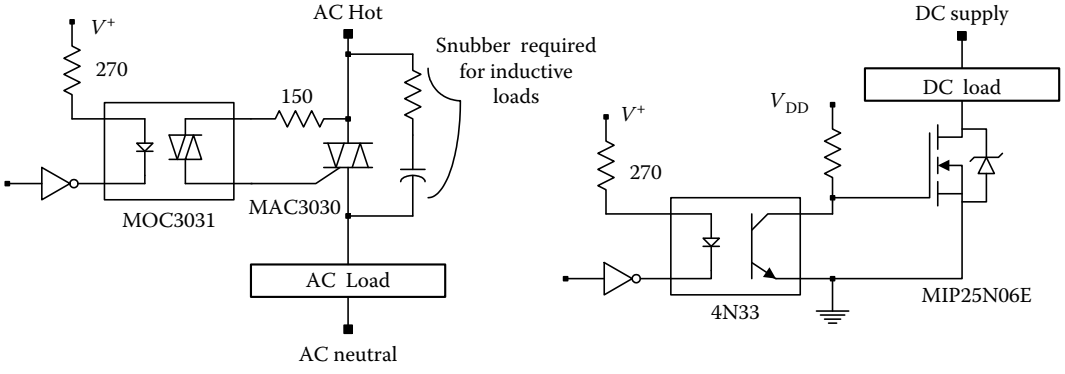


FIGURE 9.85 Circuits for opto-isolated power control.

supply to a load. With a set of bits, software can produce codes that control a device such as a printer. Since the octal latch in Figure 9.84 has a tristate output, O/E could be controlled by address and control signal decode from another microprocessor to place the octal latch output on another data bus. Thereby we accomplish data transfer between two machines. Or, software can produce binary time functions to control some process. An analog signal can be generated by outputting to a D/A converter, as shown in Figure 9.86.

Through a parallel input port a program can receive information. This may be the position of a set of switches, as shown in Figure 9.87, or the output of some kind of an encoder, such as a keypad encoder, rotating shaft position encoder, A/D converter, as shown in Figure 9.88, or even data from another microprocessor.

For data transfer from one system to another, the sender needs to know if the intended recipient is ready to receive data. And, if the data has been sent, then has it been received? To facilitate asynchronous

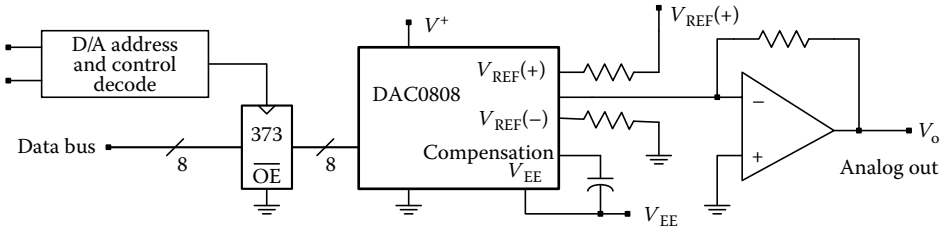


FIGURE 9.86 Output to a D/A converter.

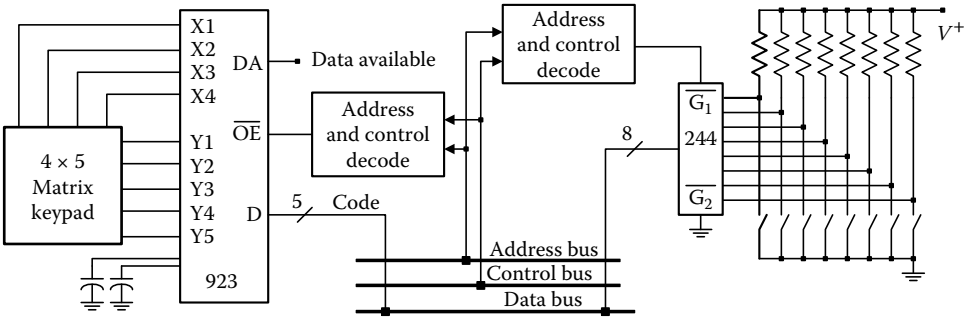


FIGURE 9.87 Circuits to input switch closures.

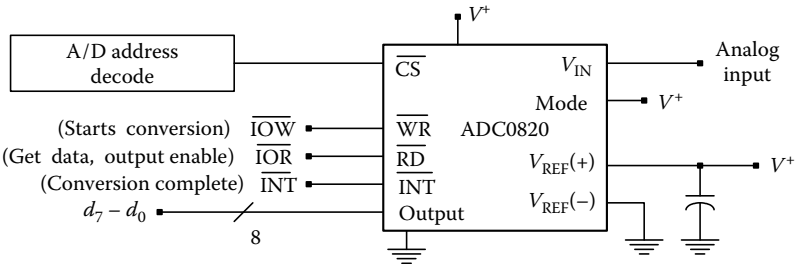


FIGURE 9.88 Write controlled A/D that includes sample and hold.

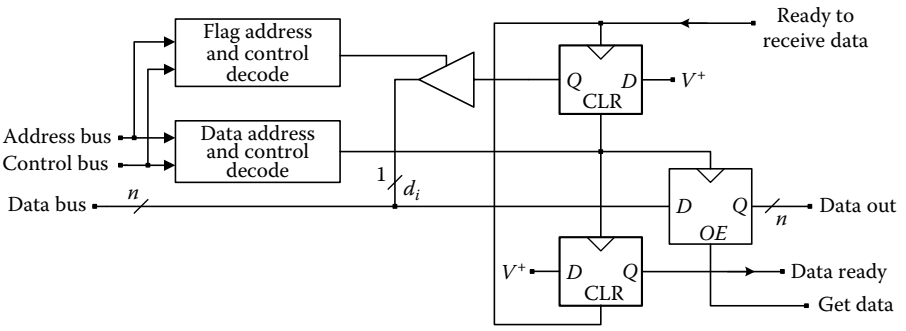


FIGURE 9.89 Output port with handshaking.

transfer of data, consider the output interface shown in Figure 9.89. This interface is the output port given in Figure 9.84 and additional circuitry to support ready to receive data and data ready flags. The sender can poll (obtain and check) the ready to receive data flag by inputting the data at the address assigned to flags. If data bit d_i is reset (logic 0), then continue polling the ready to receive data flag. If d_i is set (logic 1), then write data to the address assigned to be the data port, indivisibly reset the ready to receive data flag, and set the data ready flag. Thereby the sender cannot find that the ready to receive data flag is set until the recipient has found that the data ready flag is set, obtained the data, and set the ready to receive data flag, however little or much time this takes. Notice that setting the ready to receive data flag indivisibly resets the data ready flag. Thereby the recipient cannot find that data is available until it is new data, however little or much time this takes. There must be provision to reset the ready to receive data and data ready flags at system reset. The exchange of status information and the action among the flags is called handshaking. Figure 9.90 shows an input interface that includes handshaking. Here too, there must be provision to clear the data ready and ready to receive data flags at system reset.

To accomplish data transfer, the I/O methods of Figures 9.89 and 9.90 require that a program must continually poll the flags concerned with I/O. This works as long as the microprocessor is not required to do anything else. If the microprocessor must be used to do other tasks, then I/O can be serviced by either periodically polling I/O status flags or by using I/O status flags to interrupt the microprocessor while it is executing some program. Periodic polling of status flags may or may not be satisfactory. This depends on how much and how often I/O occurs.

Interrupt processing of the Z80, as well as the M68HC11, starts by asserting an interrupt input of the microprocessor. If more than one device must interrupt the microprocessor then using a microprocessor

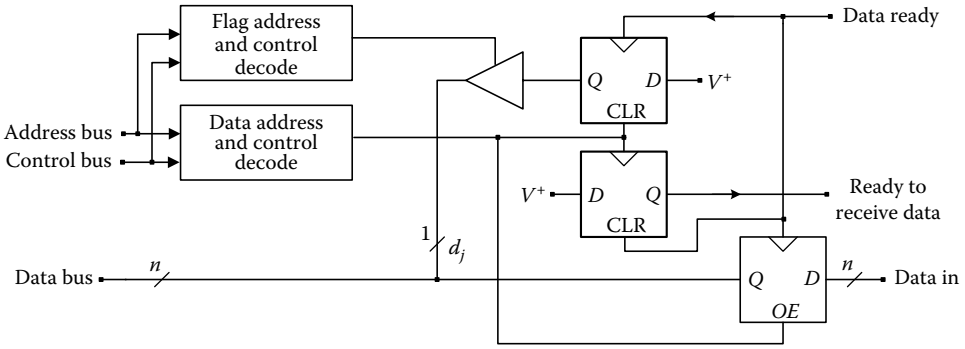


FIGURE 9.90 Input port with handshaking.

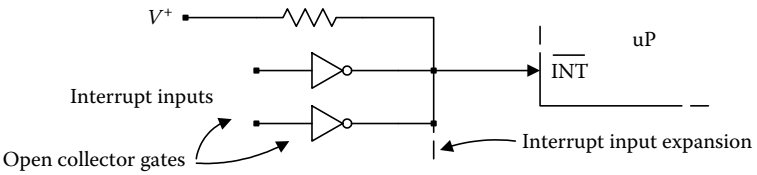


FIGURE 9.91 A wire OR'd circuit.

with many interrupt inputs might be useful. Or, all interrupt sources can be wire OR'd with open collector gates, as shown in Figure 9.91, to produce one interrupt signal. For example, the flip-flop output *Q* for the ready to receive flag in Figure 9.89 and the flip-flop output *Q* for the data ready flag in Figure 9.90 could each be connected to an inverter input in Figure 9.91. Either or both flags can then interrupt the processor. Any number of additional open collector gates could be added to include additional interrupt sources. By this method, the microprocessor discovers the occurrence of an interrupt, but must find out which device caused the interrupt.

If the Z80 is operating in direct interrupt mode, then the interrupt service routine, which starts at address 0038H, can input all of the status flags, check which flags are set, and respond with service priority determined by the software. Thus, any number of interrupts can be accommodated.

If the Z80 is operating in indirect vectored mode, then an interrupt causes it to respond with an active $\overline{\text{INTA}}$. The processor then accepts a byte from the data bus that must be provided by the interrupting device, and this byte, along with the content of the I register, is used as an address to get the interrupt service routine address from memory that is then loaded into the PC. Thus, the I register points to an interrupt vector table, and each interrupting device can point to the address of its own service routine.

If there is more than one interrupt source, then, since only one device is allowed to drive the data bus, a priority must be established, where the interrupting device with highest priority is serviced first. A daisy chain can be used so that hardware determines priority. This is shown in Figure 9.92. The device connected to the top of the chain has the highest priority. The $\overline{\text{INTA}}$ signal propagates down the chain until it is blocked by a set status flag. This condition is then used to place a response to $\overline{\text{INTA}}$ on the data bus.

In applications, such as control, it is often necessary that a program keep track of elapsed time. This can be accomplished with a circuit like the one shown in Figure 9.93. The 16 bit counter is clocked by the system clock or a clock derived from the system clock. The state of the counter is compared to the output of two registers, and a match, which clears the counter, is then used to interrupt the microprocessor. Thus, through I/O ports, software can set the time interval between interrupts, which can be disabled

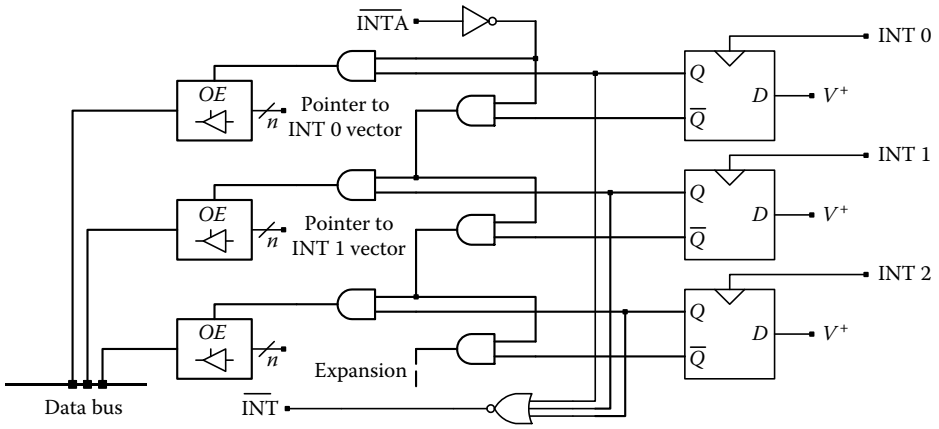


FIGURE 9.92 Priority with a daisy chain.

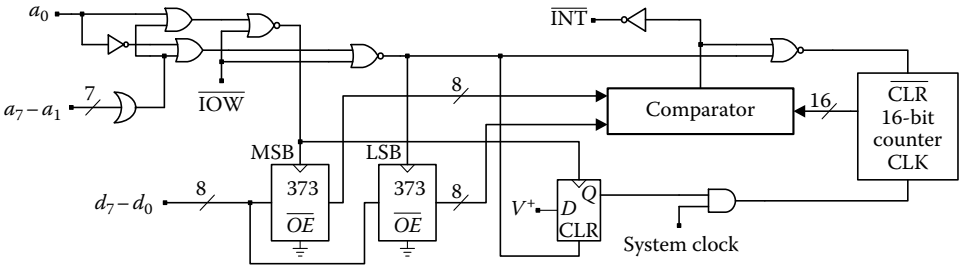


FIGURE 9.93 Timer-counter circuit.

(a write to the LSB) or enabled (a write to the MSB), and thereafter the microprocessor will be periodically interrupted to perform tasks in real time.

Solutions to interfacing problems are not unique. Variables such as the number of I/O devices, I/O device characteristics, amount of I/O, required response times, handshaking, software and hardware trade-offs, cost of parts and eventual manufacture, etc. all influence interface design. Also, there usually are numerous options to solving a particular problem. All of this requires an understanding of system (microprocessor) bus activity to properly time the occurrence of data transfer events. Then, there are issues of response to asynchronous events and status information about data transfer, polling and interrupt processing, and hardware and software trade-offs.

Solutions to common I/O problems such as parallel I/O, serial I/O, timer functions, etc. are often available within single IC packages. Typically, manufacturers of microprocessors provide a family of compatible ICs for each microprocessor. This can simplify the design task and reduce package count. Or, to reduce package count, consideration should also be given to combining gate level hardware into a single package of a PLD. Standardizing I/O can also reduce I/O coats. For example, the universal serial bus (USB) has been widely accepted to accommodate a great variety of I/O requirements.

9.4.7 Design with a Microcontroller

The M68HC11 is a microcomputer system within a single IC package. It is available in a 48 pin DIP and a 52 pin leaded chip carrier (LCC). In addition to the microprocessor, it can be configured for a variety of resources within the same package. These additional internal resources can be ROM (PROM or EPROM), RAM, parallel I/O ports, serial I/O ports, EEPROM, timer, and an A/D converter. Through a reserved bank of 64 status/control and I/O registers, software can select modes of operation of this additional hardware from a multitude of options. With these additional resources, especially the timer and A/D converter, the device is also called a microcontroller.

The programming model of the M68HC11 is shown in Figure 9.94. Registers A and B are both accumulators, and they can be used together as accumulator D for 16 bit add, subtract, multiply, and

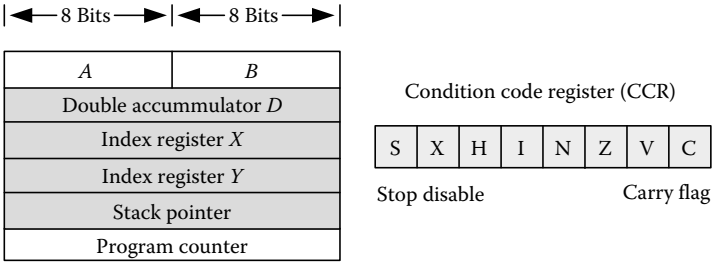


FIGURE 9.94 Programming model of the M68HC11.

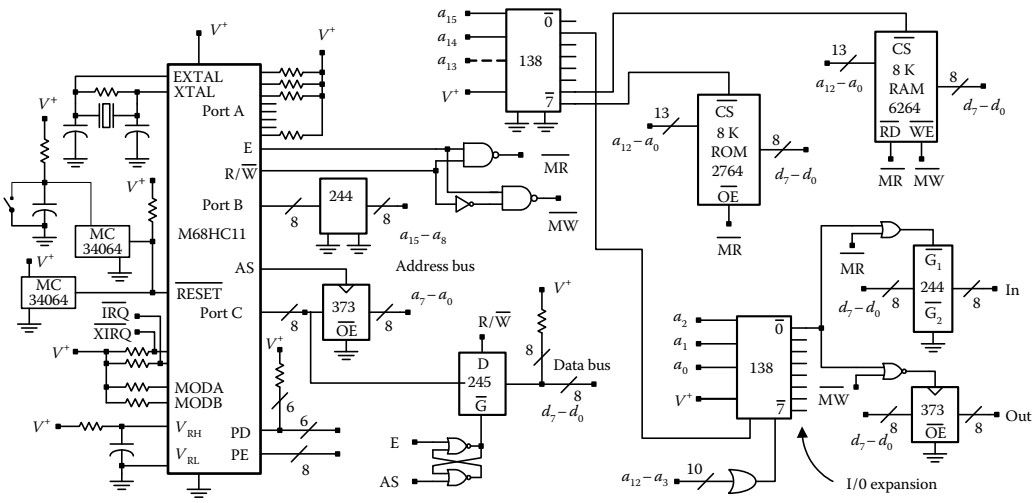


FIGURE 9.95 M68HC11A8-based microcomputer.

divide instructions. The machine language (op-codes), instruction set, and mnemonics are not the same as for the Z80, and so, the M68HC11 has its own assembly language.

The microcontroller also has different modes of operation. One of four modes is determined by the inputs at the two pins labeled MODA and MODB. These modes are MODB = 0, MODA = 0, special bootstrap, (01) special test, (10) normal single chip, and (11) normal expanded. We will consider the single chip and expanded modes of operation.

The M68HC11A8 is a version that has eight analog input channels, and in normal expanded mode of operation its pin functions are shown in Figure 9.95, which gives a diagram of a conventional microcomputer. Notice the selection of the normal expanded mode. For convenient reference, 38 of the 52 pins are grouped and labeled as follows: port A = pins 27–34, port B = pins 35–42, port C = pins 9–16, port D = pins 20–25, and port E = pins 43–50. The remaining pins are used for power supply (V_{SS} = pin 1 and V_{DD} = pin 26), the crystal for the internal clock circuit (pins 7 and 8), $\overline{\text{RESET}}$ (pin 17), nonmaskable interrupt ($\overline{\text{XIRQ}}$ = pin 18), maskable interrupt ($\overline{\text{IRQ}}$ = pin 19), control bus (AS = pin 4, E clock = pin 5, and R/$\overline{\text{W}}$ = pin 6), and analog input range (VRL = pin 51 to VRH = pin 52).

Port B provides the upper address byte. Port C is the data bus, and, to economize on the package pin count, the lower address byte is time-multiplexed with the data bus. The control signal AS, for address strobe, signifies when the data bus contains the lower address byte. The 74HC373 captures the lower address byte when AS is active. Thus, the external system address bus consists of port B and the output of the octal latch. The remainder of the control bus consists of the R/$\overline{\text{W}}$ signal and the E clock signal, which has a frequency equal to one-fourth the system clock frequency.

Since all I/O is memory mapped, the timing of system bus activity is especially straightforward. Figure 9.96 shows M68HC11 timing for all external memory references. After the lower address byte has been captured, the E clock goes high, the data bus is available for data transfer, and the R/$\overline{\text{W}}$ signal determines the direction. The E clock and the R/$\overline{\text{W}}$ signal together control read and write operations. During a read operation, the M68HC11 begins to follow the data bus at the leading edge of the E clock, and it accepts the content of the data bus at the trailing edge of the E clock. Thus, external hardware has slightly less than 2 clock cycles to provide valid data. During a write operation, the microprocessor begins to drive the data bus with valid data at the leading edge of the E clock. Thus, external hardware has slightly less than 2 clock cycles to follow the data bus, and at the trailing edge of the E clock it must accept the content of the data bus. All external memory references require one E clock cycle.

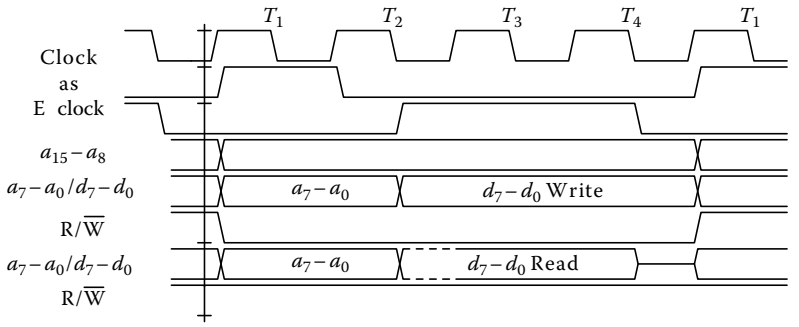


FIGURE 9.96 M68HC11 timing diagram for external memory references.

At power-up, an active low reset is accomplished with the MC34064 device in Figure 9.95. The output of this device switches low whenever its input supply is below a particular limit, and the output goes into tristate whenever its input exceeds a particular limit. Upon reset, the microprocessor initializes internal registers to default conditions, and then it reads memory locations $FFFE and $FFFF to obtain the reset vector that it loads into the PC. Thus, a designer can locate code that is to be executed at reset almost anywhere in the memory space.

Interrupt processing occurs in a similar way. Just prior to fetching each op-code the microprocessor checks the $\overline{\text{IRQ}}$ input. If it is active, and if the *I*-bit (interrupt mask) in the CCR is reset, then the microprocessor stacks 9 bytes (PC, D, IX, IY, and CCR) from the registers in its programming model, disables further interrupts by setting the *I*-bit, and then it reads memory locations $FFF2 and $FFF3 to obtain a vector that is loaded into the PC. The vector is the address of (points to) the first instruction of the $\overline{\text{IRQ}}$ interrupt service routine, which means that a designer can locate this service routine almost anywhere in the memory space.

The response to many interrupting sources can be handled in the same way as is illustrated by Figures 9.89 through 9.91, where the flags of interrupting devices are wire OR'd into one interrupt input to the microprocessor. Then, upon an interrupt, the service routine must first find out which device caused the interrupt. The control bus does not provide any special means to accomplish this. Moreover, as the service routine responds to each interrupt, there must be provision to clear each associated interrupting flag.

With the same vectored method, the M68HC11 is designed to accommodate numerous other external and internal sources of interrupts such as internal timer overflow, illegal op-code, software interrupt, nonmaskable interrupt, to name a few. Each interrupt is associated with two particular memory locations within the vector table, where its vector must be placed. This provides flexibility, and it does not require additional hardware for a reaction by an interrupting device to further control signals.

With the resources and method of interrupt processing of this microprocessor, the memory map of an M68HC11-based system is somewhat fixed. Figure 9.97 shows the memory map of the system given in Figure 9.95.

Through access to the system bus, additional memory and I/O ports can be added to this system. There are some restrictions. Internal address and control signal decoding positions the internal EEPROM block to start at memory location $B600. Because of its long write cycle time, we cannot store data into this memory with just a regular memory write instruction. There must be nonvolatile memory at the top end of the memory space to hold the vector table. The positions of the internal RAM and the control register block are programmable. The default locations are given in Figure 9.97. They can be positioned to start at any 4K boundary according to the data stored in the INIT register, which has address $103D after reset.

Many pins of the M68HC11 serve more than one purpose. For example, the mode select inputs are read during reset to determine the mode of operation. Thereafter, MODA is an active low output that

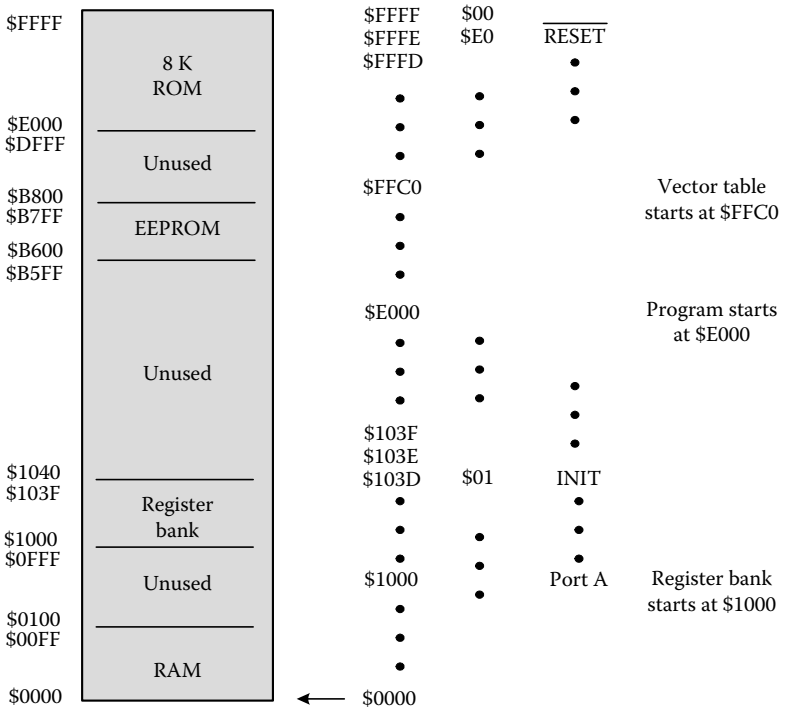


FIGURE 9.97 Memory map and sample ROM content.

becomes active during the first E cycle of each instruction. This is useful for debugging to know when the data bus must hold an op-code. The MODB pin can be used to provide standby power to maintain the content of the internal RAM when V_{DD} is not present.

Port A can serve as three input pins (PA0–PA2), four output pins (PA3–PA6), and one pin (PA7) that can be configured for input or output, depending on the data direction control bit labeled DDRA7 in the PACTL register at memory location $1026. From a programmer's viewpoint, port A is a memory register with address $1000, and the register that controls its I/O functionality is another memory register.

The M68HC11 has a free-running 16 bit counter that is clocked by the output of a programmable prescaler, which is clocked by the E clock. The 2 bytes of the counter can be read through buffers at locations $100E (high byte) and $100F (low byte). All M68HC11 timer functions are based on this counter. The 29 registers at locations $100B–$1027 are all concerned with using and configuring various kinds of timer functions. Thus, software can initiate a periodic real-time interrupt (RTI), which has an associated RTI vector located at $FFF0 and $FFF1.

Another purpose of port A is to provide access to some of these timer functions. For example, pins PA0–PA2 can be used to measure the edge-to-edge time durations of incoming pulses. The M68HC11 contains circuitry for edge detection, and for each input the edge polarity that is to be detected is programmable. Also, for each input there is a unique interrupt vector location associated with edge detection. Thus, to measure the period of a pulse train input at PA0, for example, the counter state must be captured upon an edge detection. For PA0, counter capture is controlled by the IC3F flag in register TFLG1 at location $1023. Registers at locations $101A and $101B receive the counter state. For PA0, edge detection interrupt is enabled with the IC3I flag in register TMSK1 at location $1022. The vector for this interrupt is located at $FFEA and $FFEB. If the *I*-bit in the CCR is reset, then an edge at PA0 will receive interrupt service, which should read the captured counter. Comparing counter states from successive edges of the same polarity can then be used to find the period.

Port D can serve as a general purpose 6-bit I/O port with address $1008. The direction of each bit is programmable through the DDRD register at location $1009. When serial communication is enabled, this port provides asynchronous serial input (RxD) at PD0 and serial output (TxD) at PD1. The baud rate is controlled by the contents of the BAUD register at location $102B. The five registers at locations $102B–$102F are all concerned with using and configuring asynchronous serial communication. From a programmer's viewpoint serial I/O are accomplished by a parallel read from and parallel write to memory location $102F, respectively. There are actually two responsive registers at this address that are distinguished by the R/$\overline{W}$ signal.

The other four pins of port D provide synchronous high speed serial communication. Pins PD3 and PD2 are used for transmitting and receiving serial data, respectively. Pin PD4 carries a clock signal to synchronize data transfer, and pin PD5 can be used to indicate the start of a data transfer. The three registers at locations $1028–$102A are all concerned with using and configuring synchronous serial communication. To reduce package size, pin count and communication paths, there are numerous devices that use serial I/O. Through port D, the M68HC11 can communicate with, for example, serial in LED/LCD display drivers, serial data out A/D converters, serial in/out EEPROMs, or even another microprocessor.

Port E can serve as an 8-bit digital input port with address $100A. These inputs are also each connected to a sample and hold circuit and then to an eight channel analog MUX, the output of which goes to a successive approximation A/D converter. The A/D converter produces an unsigned 8-bit number that is proportional to a DC voltage in the range VRL to VRH. An analog input equal to VRL (VRH) yields an A/D conversion result of $00 ($FF). Each A/D conversion requires 32 E clock cycles. Flags in register ADTCL at location $1030 control A/D conversion. Four consecutive conversions of a single channel, or one conversion of each of the lower four or the upper four channels can be obtained, depending flags in the ADCTL register. A/D conversion results are available from registers at locations $1031–$1034.

Needless to say, the M68HC11 is a complex machine with many options among its numerous features under software control. Features such as software security, failure detection and recovery, power-down/standby, and others have not been discussed.

In the single-chip mode as shown in Figure 9.98, ports B and C and the control signals serve other purposes. Here, port B is an 8-bit output port with address $1004. Pin 6, which is the R/$\overline{W}$ signal in the

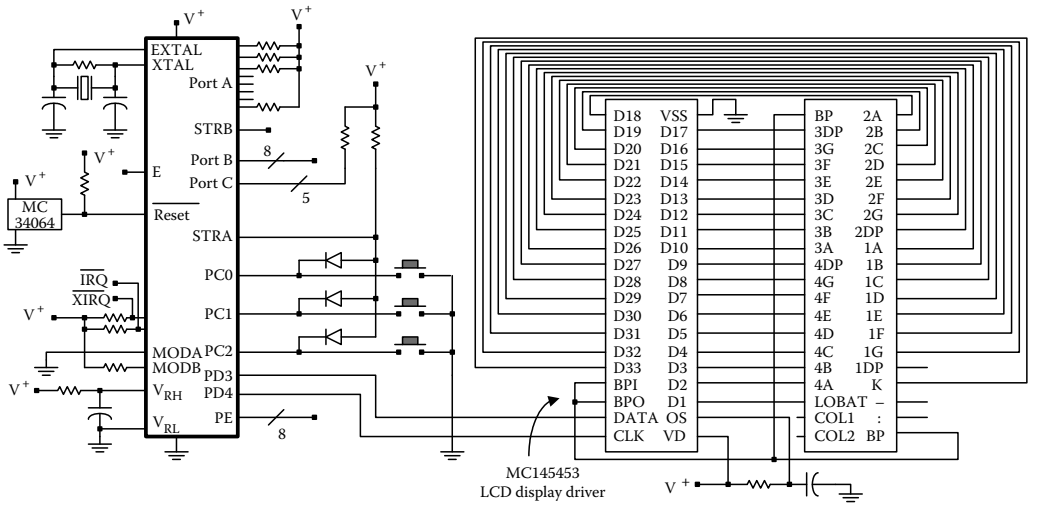


FIGURE 9.98 M68HC11 microcontroller used for a three button and four-digit LCD display device.

expanded mode, can now be configured as an output strobe (STRB) that produces a pulse whenever a write to $1004 (port B) occurs. Port C is a general purpose I/O port with address $1003, and its data direction is controlled by the DDRC register at location $1007. The eight pins of port C also go to a register with address $1005. Pin 4, which is the AS signal in the expanded mode, is now an input strobe (STRA). An edge, the polarity of which is programmable, at this input will cause the data at the port C pins to be latched into the register PORTCL with address $1005. Full handshaking is implemented, because when data are latched into the PORTCL register, the flag STAF in SR PIOC at location $1002 becomes set, and after both the flag and PORTCL have been read, the flag becomes reset.

In single-chip mode, there can be as many as 27(12) output bits and 11(26) input bits, depending on data direction control, or fewer I/O bits if some of these pins are used for A/D, serial communication, or timer functions. Moreover, software must reside in the internal 8K byte ROM of the M68HC11A8 or the internal 12K byte ROM of the M68HC11E9. Like external ROM, this ROM is positioned at the top end of the memory space. It must receive its content at the time of manufacture. With some restrictions, it can be enabled or disabled by the ROMON flag of the CONFIG register at location $103F. This is a special one byte EEPROM so that it will retain its content through power-down and power-up cycles. The M68HC711E9 is an EPROM version of the M68HC11E9. The EPROM can be programmed/erased in the field for development in the single-chip mode. There is also an OTP version of the E9. There are several other members of the M68HC11 family of microcontrollers with varying amounts of hardware resources.

Recently, the M68HC11 has evolved into the M68HC12. In addition to more of the resources available in the M68HC11, the M68HC12 has two expanded modes of operation using either an 8 bit data bus or a 16 bit data bus. Most noteworthy of this processor is its extensive and complex instruction set. In particular, there are instructions that make it convenient to implement fuzzy logic controllers, including fuzzification, inference engine, and defuzzification computing.

In contrast to the evolution of the M68HC11, the MSP430 microcontroller family from Texas Instruments has a very small instruction set, like an RISC machine, with instructions executing in single clock cycles. With its 14 bit A/D converter and fast multiply and accumulate instructions, it can be applied to do digital signal processing. Operating on 3 V, this processor is particularly well suited for low voltage, low power consumption, and portable applications.

The 8051-based family of microcontrollers from Intel has been widely utilized over the past three decades in embedded systems. It is manufactured by many companies. The Harvard architecture of its CPU is a distinctive feature of the processor.

The PIC microcontroller family from Microchip, Inc. is particularly easy to employ. Like most microcontrollers, it can be programmed in C, assembly language, and BASIC. Very low-cost development tools are available for it.

All of these microcontroller families are available with a variety of resources including ROM, RAM, EEPROM, parallel and serial I/O, A/D conversion, timers, LCD drivers, flash memory, and more within a package. When single package hardware resources can match application requirements, economical and compact hardware designs can be achieved.

9.4.8 Design Guidelines

Looking back at Figure 9.67, we see that its simplicity is deceptive. Nonetheless, in view of Figure 9.98 or even Figure 9.84, Figure 9.67 represents the typical embedded microprocessor application. To use a microprocessor requires an awareness of what may seem like an untold amount of information (facts). This should not and cannot be avoided. Time spent in the beginning to know the details will likely save time and expense that may have to be spent later to make revisions. Here, we have only raised a few issues to see the possibilities.

As we look at different processors we find that while the programming model, instruction set, and hardware resources change, hardware design is concerned with interfacing to achieve electrical, timing, and functional compatibility. There are some important principles that carry over from one

microprocessor to another. It is a component that is fundamentally intended to provide trade-offs between hardware and software. This is a matter of degree, and will vary from one application to another. Much care must be taken to suitably allocate the hardware/software trade-offs.

Software and hardware development tools such as software simulators, hardware in circuit emulators, logic analyzers, cross-assemblers and cross-compilers, real-time kernels, and others are necessary to be competitive in product development. Software will likely carry the greater burden to achieve products that function according to desired modes of operation, and it will also incur the greater development cost.

All of the manufacturers of microcontrollers mentioned above provide low-cost evaluation modules and software development tools. Once an application has been well defined, methods of implementation are sufficiently understood and required resources have been specified, it is worthwhile to exercise evaluation modules of several different microcontrollers to compare and find that microcontroller best suited for incorporation into the application.

References

1. Ayala, K. J., 2000, *The 80251 Microcontroller*, Upper Saddle River, NJ: Prentice-Hall.
2. Bierl, L., 2000, *MSP430 Family: Mixed Signal Microcontroller Application Reports*, Dallas, TX: Texas Instruments, Inc.
3. Cavenor, M. and Arnold, J., 1989, *Microcomputer Interfacing: An Experimental Approach Using the Z80*, Upper Saddle River, NJ: Prentice-Hall.
4. Cady, F. M., 1997, *Software and Hardware Engineering, Motorola M68HC12*, New York: Oxford University Press.
5. Driscoll, F. F., Coughlin, R. F., and Villanucci, R. S., 1994, *Data Acquisition and Process Control with the M68HC11 Microcontroller*, Columbus, OH: Merrill.
6. Haznedar, H., 1991, *Digital Microelectronics*, Redwood City, CA: Benjamin/Cummings.
7. Morton, T. D., 2001, *Embedded Microcontrollers*, Upper Saddle River, NJ: Prentice-Hall.
8. Noergaard, T., 2005, *Embedded Systems Architecture: A Comprehensive Guide for Engineers and Programmers*, Amsterdam/Boston: Elsevier.
9. Peatman, J. B., 1998, *Design with PIC Microcontrollers*, New York: Prentice-Hall.
10. Rafiquzzaman, M., 2005, *Fundamentals of Digital Logic and Microcomputer Design*, Hoboken, NJ: Wiley.
11. Short, K. L., 1998, *Embedded Microprocessor Systems Design, An Introduction Using the Intel 80C188EB*, New York: Prentice-Hall.
12. Spasov, P., 2002, *Microcontroller Technology, The 68HC11*, Englewood Cliffs, NJ: Prentice-Hall.
13. Uffenbeck, J., 1985, *Microcomputers and Microprocessors, The 8080, 8085, and Z80, Programming, Interfacing, and Troubleshooting*, Englewood Cliffs, NJ: Prentice-Hall.
14. Wolf, W., 2005, *Computers as Components, Principles of Embedded Computing System Design*, San Fransico, CA: Elsevier.

9.5 Systolic Arrays

Kung Yao and Flavio Lorenzelli

9.5.1 Concurrency, Parallelism, Pipelining, and Systolic Array

9.5.1.1 Motivations and Definitions

Real-time high throughput rate processing constitutes one of the most demanding aspects of modern digital signal processing. In order to achieve the desired throughput rate, various forms of concurrent operations are needed. ''Concurrency'' denotes the ability of a processing system to perform more

than one operation at a given time. Concurrency can be achieved through either parallelism or pipelining, or both. ''Parallelism'' addresses concurrency by replicating some desired processing functions many times. High throughput rate is achieved by having simultaneous operations performed by these functions on different parts of the program. On the other hand, ''pipelining'' tackles concurrency by breaking some demanding part of the task into many smaller simpler pieces, with many corresponding processing elements (PEs), so that processing can be performed in a pipeline manner. This digital pipe is arranged so that it is capable of processing the instructions and data independent of the number of PEs in the pipe. Then, high throughput rate can be achieved by having fast PEs in the pipe. As we shall see, a ''systolic array'' can exploit both the parallelism and pipelining capability of some algorithms.

The term systolic array was coined by Kung and Leiserson [2] to denote one simple class of concurrent processors, in which processed data move in a regular and periodic manner similar to that of the systolic pumping action of the blood by the heart. The earlier definition of a systolic array by Kung [3] requires (1) only a small class of PEs is in the array, with each element in a class performing identical operation; (2) all operations are performed in a synchronous manner independent of the processed data—the only control data broadcast to the PEs is the synchronous clock signal; and (3) the PEs have only nearest-neighbor communications. These regular structure and local communication properties of a systolic array are consistent with efficient modern VLSI designs. Later, various extensions of these assumptions were made (1) some of the PEs can perform a limited number of different functions, depending on the presence of some control data; (2) wavefront array allows PEs to start/end/control their own processing tasks, depending on the data; and (3) PEs can have communications to few nearby neighbors; wraparound communications among PEs located at the edge of the array are allowed.

Systolic arrays can be designed as linear arrays or two-dimensional rectangular or triangular arrays. In Figure 9.99a, consider a uniprocessor system requiring μ time unit to complete a basic operation. If some task requires N such repeated identical operations, the effective throughput rate of this system is given by $r_a = 1/N\mu$. In Figure 9.99b, consider a linear array consisting of a single pipe with N such PEs. Then the rate of this linear systolic array is given by $r_b = 1/\mu$. This demonstrates the pipelining aspects of the array. In Figure 9.99c, consider a rectangular array consisting of M pipes, with each pipe having N PEs. The rate of this rectangular systolic array is given by $r_c = M/\mu$. This demonstrates both the pipelining and parallelism of the array. While the three models in Figure 9.99 are overly simple, nevertheless they demonstrate the fact that if a given task can be designed for systolic processing, different systolic arrays can yield significantly higher throughput rates as compared to a uniprocessor of a given capability. This is the most basic aspect of systolic processing in which a higher hardware complexity is traded for a higher throughput rate.

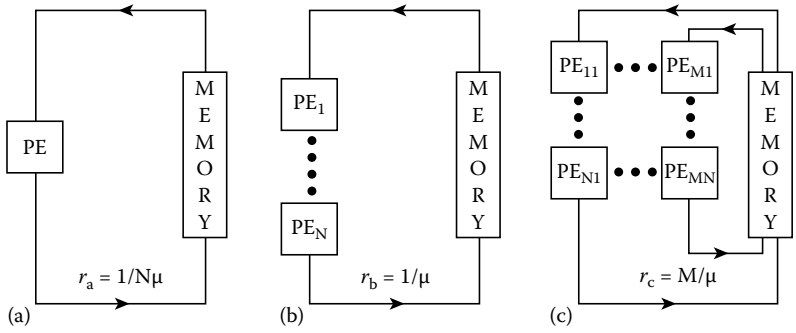


FIGURE 9.99 (a) A uniprocessor system, (b) a linear systolic array, and (c) a rectangular systolic array.

9.5.1.2 Systolic Arrays for Correlation

Consider the linear correlation of a data sequence $\{x_1, x_2, \ldots, x_M\}$ with a weight sequence $\{a_1, a_2, \ldots, a_N\}$ to yield an output sequence $\{y_1, y_2, \ldots, y_{M-N+1}\}$ given by $y_i = a_1 x_i + a_2 x_{i+1} + \cdots + a_N x_{i+N-1} = \Sigma_{J=1}^{N} a_j x_{i+j-1}$, $i = 1, 2, \ldots, M-N+1$. For the case of $N = 3$ and $M > N$, we have

$$\begin{aligned} y_1 &= a_1 x_1 + a_2 x_2 + a_3 x_3 \\ y_2 &= a_1 x_2 + a_2 x_3 + a_3 x_4 \\ y_3 &= a_1 x_3 + a_2 x_4 + a_3 x_5 \\ &\vdots \end{aligned}$$

Here, we show two of many possible systolic arrays that can implement the above correlation operations. Design B1 in Figure 9.100a uses three identical PEs to perform the accumulation (multiply and add) operation. Here, the weights a_i are preloaded to the cells and stay throughout the computation. Partial results y_i move systolically from cell to cell. Starting at the third iteration, $y_1, y_2, \ldots$, are outputted from the rightmost cell at the rate of one output per iteration. For each iteration, an x_i is broadcast to all the cells, and a y_i, initialized to zero, enters the leftmost cell. The broadcasted data x_i is marked with an arrow $\downarrow$ in Table 9.12. Indeed, by comparison we see y_1, y_2, and y_3, outputted at iteration $T = 3$, 4, and 5, agree with those given from the correlation equations.

In design B2, shown in Figure 9.100b, each input x_i is again broadcasted to each cell, each y_i stays at each cell to accumulate terms, while the weights a_i circulate around the cells in the array. A tag bit is associated with a_1 to reset the contents of the accumulator, while a tag bit is associated with a_3 to output the contents of the accumulator after the first two iterations. Data movements in design B2 are shown in

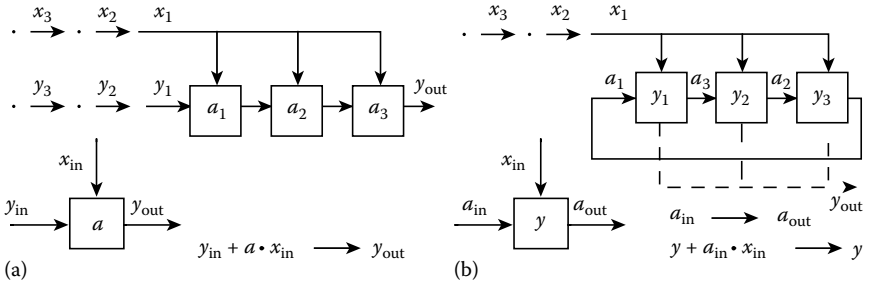


FIGURE 9.100 (a) Systolic array design B1 for correlation; and (b) systolic array B2 for correlation.

TABLE 9.12 Data Movement in Design B1

Iteration	Cell 1	Cell 2	Cell 3
$T = 1$	$\downarrow x_1\ a_1$	$\downarrow x_1\ a_2$	$\downarrow x_1\ a_3$
$T = 2$	$\downarrow x_2\ a_1$	$a_1 a_1 + \downarrow x_2\ a_2$	$x_1 a_2 + \downarrow x_2\ a_3$
$T = 3$	$\downarrow x_3\ a_1$	$x_2 a_1 + \downarrow x_3\ a_2$	$x_1 a_1 + x_2 a_2 + \downarrow x_3\ a_3 \rightarrow y_1$
$T = 4$	$\downarrow x_4\ a_1$	$x_3 a_1 + \downarrow x_4\ a_2$	$x_2 a_1 + x_3 a_2 + \downarrow x_4\ a_3 \rightarrow y_2$
$T = 5$	$\downarrow x_5\ a_1$	$x_4 a_1 + \downarrow x_5\ a_2$	$x_3 a_1 + x_4 a_2 + \downarrow x_5\ a_3 \rightarrow y_3$

TABLE 9.13 Data Movement in Design B2

Iteration	Cell 1	Cell 2	Cell 3
$T=1$	$0 + a_1 \overset{\downarrow}{x_1}$	$a_3 \overset{\downarrow}{x_1}$	$a_2 \overset{\downarrow}{x_1}$
$T=2$	$a_1x_1 + a_2 \overset{\downarrow}{x_2}$	$0 + a_1 \overset{\downarrow}{x_2}$	$a_2x_1 + a_3 \overset{\downarrow}{x_2}$
$T=3$	$a_1x_1 + a_2x_2 + a_3 \overset{\downarrow}{x_3} \rightarrow y_1$	$a_1x_2 + a_2 \overset{\downarrow}{x_3}$	$0 + a_1 \overset{\downarrow}{x_3}$
$T=4$	$0 + a_1 \overset{\downarrow}{x_4}$	$a_1x_2 + a_2x_3 + a_3 \overset{\downarrow}{x_4} \rightarrow y_2$	$a_1x_3 + a_2 \overset{\downarrow}{x_4}$
$T=5$	$a_1x_4 + a_2 \overset{\downarrow}{x_5}$	$0 + a_1 \overset{\downarrow}{x_5}$	$a_1x_3 + a_2x_4 + a_3 \overset{\downarrow}{x_5} \rightarrow y_3$

Table 9.13. Note that resets occur at cell 1 at iteration 1, cell 2 at iteration 2, cell 3 at iteration 3, cell 4 at iteration 4, etc. Similarly, output y_1 occurs from cell 1 at iteration 3, y_2 from cell 2 at iteration 4, y_3 from cell 3 at iteration 5, etc.

9.5.1.3 Systolic Array Design Techniques

Systolic array designs, as shown above for the correlation case, can be obtained by ad hoc approaches. More formal procedures for the systematic design of systolic arrays have been proposed by Moldovan [7], Quinton [9], Kung [5], Rao [10], Darte and Delosme [1], and others. All those more formal procedures are collectively referred to as dependence graph mapping techniques for systolic array design.

In this approach, an algorithm must be formulated in the ''single assignment algorithm'' form. Each variable has a unique value during the evaluation of the algorithm. Those variables with multiple values can be converted to single values by vectorizing the variables through the introduction of new indices. As an example, consider the matrix multiplication of $C = AB$, where $A = [a_{ik}]$ is $N_1 \times N_3$, $B = [b_{kj}]$ is $N_3 \times N_2$, and $C = [c_{ij}]$ is $N_1 \times N_2$. A conventional formulation of this algorithm contains the expression, $c_{ij} = c_{ij} + a_{ik}\, b_{kj}$ for $i = 1$ to N_1, $j = 1$ to N_2, and $k = 1$ to N_3. We note, c_{ij} has multiple values for $k = 1, \ldots, N_3$. We can modify it to have single values by replacing it by the variable c_{ijk}. The previous equation for c_{ij} then becomes $c_{ijk} = c_{ij(k-1)} + a_{ik}\, b_{kj}$, $c_{ij0} = 0$, c_{ijN_3}, $i = 1, \ldots, N_1$, $j = 1, \ldots, N_2$, $k = 1, \ldots, N_3$.

All algorithm variables are assumed to be indexed variables with V variable names, denoted by the generic names of V_m, $1 \leq m \leq V$. In the above matrix–matrix multiplication problem $V = 3$, and we can take $X_1, = c$, $X_2 = a$, and $X_3 = b$. For each variable name, the domain of the index vectors is a subset in an S-dimensional space. This subset is called the algorithm's ''index space'' and S is its dimension. For most iterative signal processing problems, time is usually one of the index space coordinates. For the preceding matrix–multiplication problem, we need to propagate a_{ik} across the j variables as well as b_{kj} over the i variables in order to perform the basic multiplication operation. These and a_{ik} and b_{kj} are propagating variables because they involve no computations, but need to be made available at various stages of the computation. In the matrix–matrix problem, clearly $S = 3$ and the index space is $\mathbf{S}_0 = \{(i, j, k)\colon 1 \leq i \leq N_1, 1 \leq j \leq N_2, 1 \leq k \leq N_3,\}$. Furthermore, the initializations of the new variables are given by $a(i, 0, k) = a_{ik}$, $b(0, j, k) = b_{jk}$, $c(i, j, 0) = 0$, $c(i, j, N_3) = c_{ij}$, and the algorithm is finally given by $a(i, j, k) = a(i, j, -1, k)$, $b(i, j, k) = a(i - 1, j, k)$, $c(i, j, k) = c(i, j, k - 1) + a(i, j, j)b(i, j, k)$, for $(i, j, k) \in \mathbf{S}_0$.

In general, a point (or node) in the index space is called an ''index point.'' Thus, $X_m(I)$ is the variable X_m defined at the index point I. A dependence graph mapping is a representation of a single assignment algorithm, where the dependencies among the variables are represented by directed arcs among the nodes. A basic property of the class of algorithms of interest is that of ''shift-invariance.''

An algorithm is shift-invariant if the dependence graph is regular. That is, $X(I)$ depends on $Y(J)$, then $X(I + K)$ depends on $Y(J + K)$ for all I, J, and K in the index space. Three well-known shift-invariance algorithms include

1. Uniform recurrence equations (URE): $X_1(I) = F_1(X_1(I - D_1), \ldots, X_V(I - D_V))$, $X_i(I) = X_i(I - D_i)$, $2 \leq i \leq V$. Computation occurs only in $F_1(\cdot)$ and propagations in all the other variables, Clearly, the final form of the above matrix–matrix multiplication algorithm is a URE algorithm with

$V = 3$, $F_1(\cdot) = c(i, j, k)$, with $X_1(\cdot) = c(\cdot)$, $X_2(\cdot) = a(\cdot)$, $X_3(\cdot) = b(\cdot)$, $I = (i, \mathrm{i}, k)$, $D_1 = [0, 0, 1]^{\mathrm{T}}$, $D_2 = [0, 1, 0]^{\mathrm{T}}$, and $D_3 = [1, 0, 0]^{\mathrm{T}}$.

2. Generalized uniform recurrence equations (GURE): $X_m(I) = F_m(X_{m1}(I - D_{m1}), \ldots, X_{mk(m)}(I - D_{mk(m)}))$, $1 \leq m \leq V$, where $m_1, \ldots, m_{k(m)}$, belongs to $\{1, \ldots, m\}$. In GURE, we can have computations in all V functions of $F_m(\cdot)$. The number of independent variables $m_{k(m)}$ depends on each m. The shift index dependence, $I - D_{mi}$, is fixed for each X_{mi}.
3. Regular iterative algorithm (RIA): $X_m(I) = F_m(X_{m1}(I - D_{m1,m}), \ldots, X_{mk(m)}(I - D_{mk(m)m}))$, $1 \leq m \leq V$. Here, the shift index dependency, $I - D_{m_i m}$, is not fixed but is a function of m_i and m.

Each processor of the systolic array is assumed to have all the necessary computational modules to compute $F_m(\cdot)$. For URE, we need only one such module, but for GURE and RIA, we need V modules. The time required for the computation of $F_m(\cdot)$ is denoted by τ_m, and the minimum time between such computations is denoted by h_m. In most cases, we can set τ_m and h_m to unity. The design of a processor array to perform the algorithm requires spatial and temporal assignments. Each $X_m(I)$ must be assigned to a processor at each integral time slot. The processor "allocation function," $A(I)$ assigns all variables with the index I to the processors in the array. The "scheduling function," $S_m(I)$, assigns the start of the computation for the variable $X_m(I)$. The simplest form of scheduling and processor allocations are based on the projection of the high multidimensional dependence graph onto the lower dimensional processor array. Variables represented by nodes in the dependence graph are mapped to processors which perform the computations. The directed arcs of the dependence graphs are transformed to physical communication links in the processor array.

The essence of the allocation function $A(\cdot)$ is thus to return for every index value $I \in \mathbf{S}_0$ a vector which indicates the processor in charge of the computation represented by a point in a lower dimensional space. Analogously, the scheduling function $S_m(\cdot)$ provides the relative start of the execution for the computation indexed by I. These two functions cannot be chosen independently because two computations assigned to the same processor cannot be scheduled for the same time ("compatibility constraint"). Additional details on this constraint are given later. While in principle $A(\cdot)$ and $S_m(\cdot)$ can be any function, we shall consider only "affine functions," in the sense that $A(I) = A^{\mathrm{T}}I$, $S_m(I) = \lambda^{\mathrm{T}} i + \gamma_m$, where A is a suitable matrix, λ a vector, and γ_m an integral constant.

The dependence graph of an algorithm can be interpreted as a "lattice" embedded in a multidimensional integral space (i.e., a proper bounded subset of $\mathbf{Z}^s$, where $\mathbf{Z}$ is the set of relative integers), enclosed in a convex polyhedron. We assume the lattice to be "dense" in the sense that all the integral points in it correspond to actual computations. The whole procedure of mapping an algorithm onto a systolic-type processor consists of two conceptually different but interdependent operations of using a space transformation and a time transformation. The former actually "projects" the dependence graph onto a lower dimensional structure which then can be mapped one-to-one onto the physical array, while the latter gives the start of the execution of each computation.

For simplicity, consider the projection of the S-dimensional space onto an $(S-1)$ dimensional processor space. The more general problem of projecting the dependence graph onto an $(S-p)$-dimensional space ($p \geq 1$) can be expressed using a similar but more involved notation and is omitted here. Instead of considering allocation functions, we refer to the "projection vector u," which is orthogonal to the processor space onto which we project. Assume that we have chosen both the projection and the scheduling vectors (u and λ, respectively). For normalization purposes, they are chosen to be coprime vectors, such that the greatest common divisor of their components is 1, and their first nonzero element is positive.

Two sets of constraints must be satisfied by u and λ. Assume nodes I and J are located along a direction parallel to the projection vector u such that $J = I + \alpha u$, $\alpha \in \mathbf{Z}$. Then, the computations associated with the two nodes will be projected onto the same processor. Consequently, compatibility constraint requires that they be performed at different times. Analytically, this is equivalent to $|\lambda^{\mathrm{T}} u|\ \lambda \max_{m=1,\ldots,V} h_m$, which for $h_m = 1$ simplifies to $|\lambda^{\mathrm{T}} u| > 0$. Thus, for this case λ and u cannot be perpendicular. Furthermore, the

quantity $c = \Delta|\lambda^T u|$ represents the number of time slots between successive calculations scheduled on the same processor. $1/c$ is sometimes called the "efficiency" of the processors because the larger the c, the more time the processors can idle. One common approach is to select the projection vector and the scheduling vector to achieve the highest efficiency, with c being as close to 1 as possible.

Consider the case in which the variable $X_m(I)$ depends on $X_n(I - D_{nm})$. The "precedence constraint" implies the calculation of $X_n(I - D_{nm})$ must be scheduled to be completed before the start of the calculation of $X_m(I)$. Analytically, the precedence constraint is equivalent to $\lambda^T D_{nm} + \gamma_m - \gamma_n \geq \tau_n = 1$, for all $1 \leq m \leq v$ and for all dependences D_{nm}. If the γ constants are chosen to be all equal, the precedence constraint becomes $\lambda^T D_{nm} \geq 1 \; \forall \, m = 1, \ldots, V$.

Assume the precedence and compatibility constraints are satisfied and λ and u are coprime vectors. Then it is possible to extend both vectors to two unimodular matrices. A matrix with integral entries is called "unimodular" when its determinant is equal to ± 1. This implies that they admit integral inverses. The unimodular extension of coprime vectors is not unique. We will choose U and Λ to be the unimodular extended matrices that have u and λ, respectively, as their first columns. It is possible to show that the columns of any S-dimensional unimodular matrix can constitute a basis for the space Z^S. Moreover, if we denote $\sigma_1, \ldots, \sigma_s$ to be the columns of $\Sigma = U^{-T}$, then we have $\sigma_1^T u = 1$ and $\sigma_1^T u = 0$ for all $i = 2, \ldots, S$. Therefore, $\{\sigma_2, \ldots, \sigma_s\}$ will be a basis of the processor space of the resulting logic array. Similarly, the first column of T (the inverse of Λ^T) t_1, represents the direction in which time increases by one step; i.e., it is the vector defining the hyperplane of the points computed at the same time. The other columns of T (denoted by $t_2, \ldots, t_s$) are a basis of such a hyperplane.

If we denote by $\Sigma_+ = [\sigma_2, \ldots, \sigma_s]$ the matrix basis of the processor space, the allocation function and the scheduling function have the form $A(I) = \Sigma_+^T I, S_m(I) = \lambda^T I + \gamma_m$, $m = 1, \ldots, V$. With these elements we can have the complete description of the final array. The processors are labeled by $A(I) = \Sigma_+^T I$ as I ranges over the index space. The dependences D_{nm} are mapped onto communication links $\Sigma_+^T D_{nm}$ and the delay registers on such links must be in number equal to $\lambda^T D_{nm} + \gamma_m - \gamma_n - \tau_n = \lambda^T D_{nm} - 1$.

Reconsider the systolic correlation problem using the weights $\{a_1, \ldots, a_k\}$ and the data $\{x_1, \ldots, x_n\}$, as discussed earlier. Recall the correlation is given by $y_i = a_1 x_i + a_2 x_{i+1} + \cdots + a_k x_{i+k-1}$, $1 \leq i \leq n + 1 - k$. A recurrence equation formulation of this equation is given by $y(i,j) = y(i,j-1) + w(i,j)x(i,j), y(i,0) = 0, y_i = y(i,k)$; $w(i,j) = w(i-1,j)$, $w(1,j) = a_j$; and $x(i,j) = x(i+1,j-1)$, $x(i,0) = x_{i-1}$, all with $1 \leq i \leq n + 1 - k$, $1 \leq j \leq k$. A dependence graphical representation of these equations is shown in Figure 9.101a.

A URE reformulation of the recurrence equations yields $X_1(I) = y(i,j) = F_1(X_1(I - D_1), X_2(U - D_2), X_3(I - D_3))$, $X_2(I) = w(i,j) = X_2(I - D_2)$, $X_3(I) = x(i,j) = X_3(I - D_3)$, with the index point $I = [i,j]^T$ and

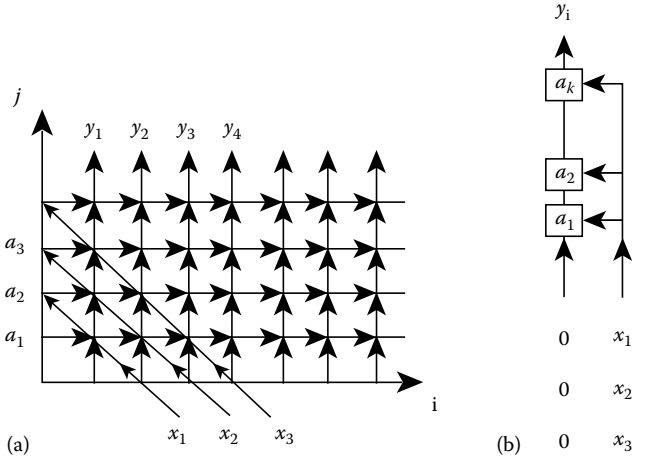


FIGURE 9.101 (a) Two-dimensional dependence graph and (b) one-dimensional dependence graph.

displacement vectors $D_1 = [0, 1]^T$, $D_2 = [1, 0]^T$, and $D_3 = [-1, 1]^T$. In particular, consider the URE representation of the B1 design based on the choice of $u = [1, 0]^T$, $\Sigma_+ = [0, 1]^T$, $\lambda = [1, 1]^T$, $\gamma_m = 0$, $m = 1$, 2, and 3, and $\tau_n = h_n = 1$, $n = 1$, 2, and 3. Then, the two-dimensional graph of Figure 9.101a is projected onto the one-dimensional graph of Figure 9.101b. Specifically, for any index $I = [i, j]^T$, the processor allocation function yields $A(I) = \Sigma_+^T I = j$, $i \le j \le k$, which is a valid projection from two-dimensions to one. On the other hand, the index point for each input data x_1 (with variable name of X_3) is given by $I = [i, l - i + 1]^T$, $l = 1, \ldots, n$. Then the scheduling function $S_3(I)$ is given by $S_3(I) = \lambda^T I = [1, 1]\ [i, 1 - i + 1]^T = i + l - i + 1 = l + 1$, $l = 1, \ldots, n$. This indicates each x_1 for the previously given I must be available at all the processors at time $l + 1$. Thus, there is no propagation and this means all the x_1 must be broadcasted to all the processors. However, the simplistic definition of a systolic design does not allow broadcasting. Indeed, the precedence constraint is *not* satisfied with D_3. That is, $\lambda^T D_3 = [1, 1]\ [-1, 1]^T = 0 \not\ge \tau_n = 1$. Of course, other choices of λ and u generate other forms of systolic array architecture for correlation. For more complicated signal processing tasks such as QR decomposition (QRD), recursive least-squares (LS) estimation, singular value decomposition (SVD), Kalman filtering (KF), etc., the design of efficient systolic array architectures are generally difficult. The dependence graph mapping technique provides a systematic approach to such designs by providing the proper selections of these λ and u vectors.

9.5.2 Digital Filters

The application of digital filtering has spread tremendously in recent years to numerous fields, such as signal processing, digital communications, image processing, and radar processing. It is well known that the sampling rate, which is closely related to the system clock, must be higher than the Nyquist frequency of the signals of interest. It follows that in order to perform real-time filtering operations when high frequency signals are involved, high-speed computing hardware is necessary.

Pipelining techniques have been widely used to increase the throughput of synchronous hardware implementations of a transfer function of a particular algorithm. Most algorithms can be described in a number of different ways, and each of these descriptions can be mapped onto a set of different concurrent architectures. Different descriptions may lead to realizations with entirely different properties, and can have a dramatic impact on the ultimate performance of the hardware implementation. Pipelining can also be used for other than throughput increase. For a fixed sample rate, a pipelined circuit is characterized by a lower power consumption. This is due to the fact that in a pipelined system capacitances can be charged and discharged with a lower power supply. Because the dissipated power depends quadratically on the voltage supply, the power consumption can be reduced accordingly.

An increase in the speed of the algorithm also can be achieved by using parallelism. By replicating a portion of the hardware architecture, similar or identical operations can be performed by two or more concurrent circuits, and an intelligent use of this hardware redundancy can result in a net throughput increase, at the expense of area. Note that VLSI technologies favor the design in which individual sections of the layout are replicated numerous times. A regular and modular design can be achieved at relatively low costs. For a fixed sample rate, parallelism can be exploited for a low power design due to the reduced speed requirements on each separate portion of the circuit.

Much work has been done in the field of systolic synthesis of finite and infinite impulse response (FIR/IIR) filters, as can be seen from the literature references. In the following subsection, we consider possible strategies that can be used to increase the throughput of the concurrent architectures of the FIR and IIR filters.

9.5.2.1 FIR Filters

FIR filters have been largely employed because of certain desirable properties. In particular, they are always stable, and causal FIR filters can posses linear phase. A large number of algorithms have been devised for the efficient implementation of FIR filters, which minimize the number of multipliers, the

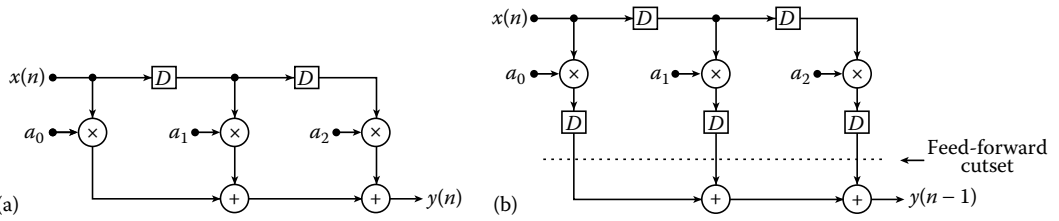


FIGURE 9.102 A three-tap FIR filter: (a) with no pipelining, the throughput is limited by the rate of one multiplication and two additions and (b) with pipelined circuit, the throughput is increased to the rate of one multiplication or two additions.

round-off noise, or the coefficient sensitivity. The generic expression that relates the output $y(n)$ at time n to the inputs $x(n-i)$ at times $(n-i)$ $i=0,1,\ldots,q$ is given by

$$y(n) = \sum_{i=0}^{q} a_i x(n-i)$$

where the $\{a_i\}_{i=0}^{q}$ are the FIR filter coefficients. Here, we consider only issues of pipelining and parallelism.

The pipeline rate, or throughput, of implemented nonrecursive algorithms such as FIR filters can be increased without changing the overall transfer function of the algorithms by means of a relatively simple modification of the internal structure of the algorithm. In particular, one set of latches and storage buffers can be inserted across any feed-forward cutset of the data flow graph. Figure 9.102b, illustrates the increase of throughout achieved by pipelining in a second-order three-tap FIR filter. The sample rate of the circuit of Figure 9.102a is limited by the throughput of one multiplication ''and'' two additions. After placing the latches at the locations shown in Figure 9.102b, the throughput can be increased to the rate of one multiplication ''or'' two additions. Pipelining can be used to increase the sample rate in all the cases in which no feedback loops are present. The drawbacks of pipelining are an increased latency and a larger number of latches and buffers.

Parallelism can be used to increase the speed of an FIR filter. Consider Figure 9.103, in which the three-tap FIR filter of Figure 9.102 was duplicated. Because at each time instant two input samples are processed and two samples are output, the effective throughput rate is exactly doubled.

As can be seen from Figure 9.103. parallelism leads to speed increase at a considerable hardware cost. For many practical implementations, parallelism and pipelining can be used concomitantly, when either method alone would be insufficient or limited by technology such as I–O, clock rate, etc.

9.5.2.2 IIR Filters

These are recursive filters in the sense that their output is function of current inputs as well as past outputs. The general I–O relationship is expressed by

$$y(n) = \sum_{j=1}^{p} a_i y(n-j) + \sum_{i=o}^{q} b_i x(n-i) \tag{9.22}$$

where the $\{a_j\}_{j=1}^{p}$ are the coefficients associated to the recursive part, and the $\{b_i\}_{i=0}^{q}$ are the coefficients associated to the nonrecursive portion of the filter. The associated transfer function is written as the following z-transform

$$H(z) = \frac{\sum_{i=0}^{q} b_i z^{-i}}{1 - \sum_{j=1}^{p} a_j z^{z-j}}$$

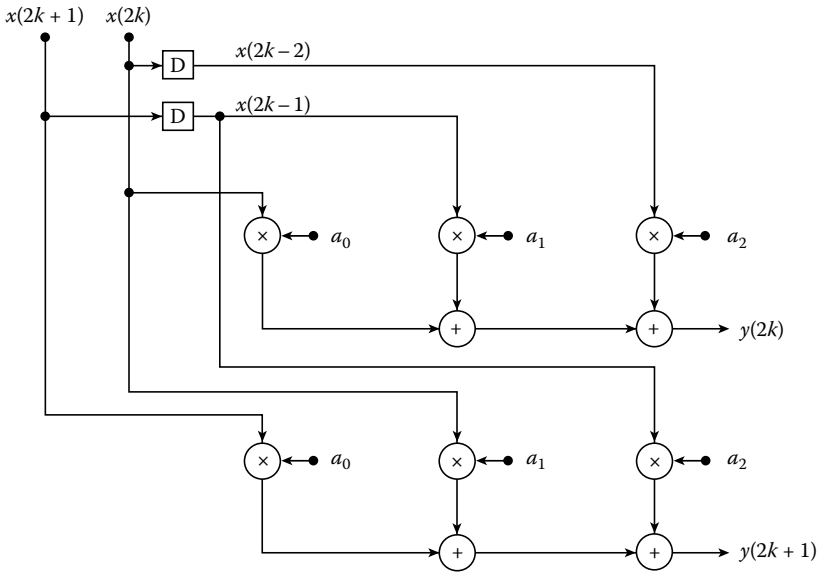


FIGURE 9.103 Three-tap FIR filter whose hardware has been duplicated to achieve double throughput rate.

For stability reasons, it is required that all the poles (i.e., the zeroes of the denominator of $H(z)$) be inside the unit circle in the z-plane.

Consider a circuit in which L loops are present each with latency τ_k, $k = 1, \ldots, L$. The number of latches present in each loop is equal to ν_k, $k = 1, \ldots, L$. Then the throughput period cannot be shorter than

$$T_{\max} \equiv \max_{k=l,\ldots,L} \left\lceil \frac{\tau_k}{\nu_k} \right\rceil$$

The pipeline can be increased by increasing the number of latches internal to the feedback loops. The computational latency associated with the internal feedback prevents one from introducing pipeline simply by inserting latches on feedforward cutsets. In fact, inserting latches in the loop would change the overall transfer function. This difficulty can be overcome by recasting the algorithm into an equivalent formulation from an I–O point of view. The transformations applied to the algorithm, prior to the mapping, have the purpose of creating additional concurrency, thereby increasing the achievable throughput rate. Without ever changing the algorithm's transfer function, additional delays are introduced inside the recursive loop. These delays are subsequently used for pipelining. In the sequel, we briefly describe two types of look-ahead techniques that generate the desired algorithmic transformations, namely the clustered and the scattered look-ahead techniques proposed by Loomis and Sinha [18] and Parhi and Messerschmitt [19], respectively. Look-ahead techniques are based on successive iterations of the basic recursion, in order to generate the desired level of concurrency. The implementation is then based on the iterated version of the algorithm.

Clustered look-ahead: In a pth order recursive system, the output at time n is a function of the past output samples $y(n-1)$, $y(n-2), \ldots, y(n-p)$. In the clustered look-ahead technique, the recursion is iterated m times so that the current output is a function of the cluster of p consecutive samples $y(n-m)$, $y(n-m-1), \ldots, y(n-m-p)$. The original order-$p$ recursive filter is emulated by a $(p+m)$th filter, where m canceling poles and zeroes have been added. In this way the m delays generated inside the feedback loop can be used to pipeline by m stages.

By iterating 9.22 m times, we can derive the following I–O relationship

$$y(n) = \sum_{j=1}^{p-1} \left[\sum_{k=j+1}^{p} a_k r_{j+m-k} \right] y(n-j-m) + \sum_{j=0}^{m-1} \sum_{k=0}^{q} b_k x(n-k-j)$$

where the coefficients $\{r_i\}$ can be precomputed off-line, and are such are such that $r_i = \Sigma_{k=1}^{p} a_k r_{i-k}$, $i > 0$, $r_0 = 1$, and $r_i = 0$, $i = -(p-1), \ldots, -1$. This implementation requires $(p+m)$ multiplications for the nonrecursive part, and p for the recursive part, for a total of $(2\,p+m)$, which grows linearly with m. The transfer function is equal to

$$H(z) = \frac{\sum_{j=0}^{m-1} \sum_{k=0}^{q} b_k z^{-k-j}}{1 - \sum_{j=1}^{p-1} \left[\sum_{k=j+1}^{p} a_k r_{j+m-k} \right] z^{-j-m}}$$

The clustered look-ahead technique does not guarantee that the resulting filter is stable because it may introduce poles outside the unit circle.

Consider the following simple example with a stable transfer function:

$$H(z) = \frac{1}{1 - 1.3z^{-1} + 0.35z^{-2}}$$

with poles at $z = 0.7$ and $z = 0.5$. The two-stage equivalent filter can be obtained by introducing the canceling pole-zero pair at $z = -1.3$, as follows:

$$H(z) = \frac{1 + 1.3z^{-1}}{(1 - 1.3z^{-1} + 0.35z^{-2})(1 + 1.3z^{-1})} = \frac{1 + 0.9z^{-1}}{1 - 1.34z^{-2} + 0.455z^{-3}}$$

Because a pole is found at $z = -1.3$, this transfer function is clearly unstable.

Scattered look-ahead: In the scattered look-ahead technique, the current output sample $y(n)$, is expressed in terms of the (scattered) p past outputs $y(n-m)$, $y(n-2m), \ldots, y(n-mp)$. The original order-p filter is now emulated by an order-mp filter. For each pole of the original filter, $(m-1)$ canceling pole-zero pairs are introduced at the same distance from the origin as the original pole. Thus, stability is always assured. The price we must pay is higher complexity, on the order of mp. To best describe the technique, it is convenient to write the transfer function $H(z)$ as a ratio of polynomials, i.e., $H(z) = N(z)/D(z)$. The transformation can be written as follows:

$$H(z) = \frac{N(z)}{D(z)} = \frac{N(z) \prod_{k=1}^{m-1} D\left(z e^{j(2\pi k/m)}\right)}{\prod_{k=0}^{m-1} D(z e^{j(2\pi k/m)})}$$

Note that the transformed denominator is now a function of z^{-m}.

Consider the example of the previous section. For the scattered look-ahead technique, it is necessary to introduce pole-zero pairs at $z = 0.7\ e^{\pm j(2\pi/3)}$ and $z = 0.5\ e^{\pm j(2\pi/3)}$. The transformed denominator equals $1 - 0.125z^{-3}$.

The complexity of the nonrecursive part of the transformed filter is $(pm+1)$ multiplications, while the recursive part requires p multiplications, for a total of $(pm+p+1)$ pipelined multiplications. Although the complexity is still linear in m, it is much higher than in the clustered look-ahead technique for a large value of p.

Parhi and Messerschmitt [19] presented a technique to reduce the complexity of the nonrecursive portion down to $O(p \log_2 m)$, applicable when m is a power of 2. This technique can be described as follows. Assume that the original recursive portion of the given IIR filter is given by

$$H(z) = \frac{1}{1 - \sum_{j=1}^{p} a_j^{(1)} z^{-j}}$$

An equivalent two-stage implementation of the same filter can be obtained by multiplying numerator and denominator by the polynomial $(1 - \Sigma_{j=1}^{p}(-1)^j a_j^{(1)} z^{-j})$, which is given by

$$H(z) = \frac{1 - \sum_{j=1}^{p} (-1)^j a_j^{(1)} z^{-j}}{1 - \sum_{j=1}^{p} a_j^{(2)} z^{-j}}$$

where the set of coefficient $\{a_j^{(2)}\}_{j=1}^{p}$ is obtained from the original set $\{a_j^{(1)}\}_{j=1}^{p}$ by algebraic manipulation. By repeating this process $\log_2 m$ times one can obtain an m-stage pipelined implementation, equivalent to the original filter. In this way the hardware complexity only grows logarithmically with the number of pipelining stages.

Bidirectional systolic arrays for IIR filtering: Lei and Yao [16] showed that many IIR filter structures can be considered as special cases of a general class of systolizable filters, as shown in Figure 9.104. These filters can be pipelined by rescaling the time so that $z' = z^{1/2}$, and by applying a cutset transformation. This time rescaling causes the hardware factorization to reduce to merely 50%, which is quite inefficient. Lei and Yao [17] later proposed two techniques to improve the efficiency of these bidirectional IIR filters.

In the first method (''overlapped subfilter scheme''), one makes use of the possibility to factor the numerator and the denominator of the given transfer function. For instance, if

$$H(z) = \frac{N(z)}{D(z)} = \underbrace{\frac{N_a(z)}{D_a(z)}}_{H_a(z)} \cdot \underbrace{\frac{N_b(z)}{D_b(z)}}_{H_b(z)}$$

where $a + b = p$, $a - b = 0$, 1, or 2, and p is the number of modules of the original transfer function. Then the two subfilters, $H_a(z)$ and $H_b(z)$, can be realized on the same systolic array of $a + 1$ modules, as in Figure 9.105. A multiplexer at the input of the array chooses the incoming data at even time instants, and the data from the output of the first module at odd time instants. The modules alternately perform operations associated to $H_a(z)$ and $H_b(z)$ in such away as to interleave the operations and have an overall 100% efficiency.

In the second technique (''systolic ring scheme'') the number of modules is about half of the order of the original transfer function. The modules of the new structure are arranged as a systolic ring, as in Figure 9.106. For example, a five-module ring can be used to implement a ten-module IIR filter: module i performs the operations associated to modules i and $(5 + i)$ of the original array, for $i = 1, \ldots, 5$. Note that in the original structure every other module is idle. The resulting ring is therefore 100% efficient.

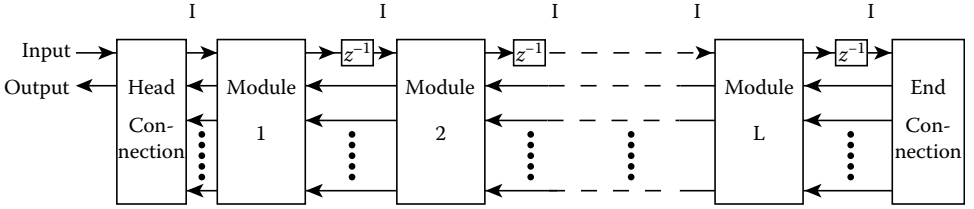


FIGURE 9.104 A general structure of bidirectional IIR filters.

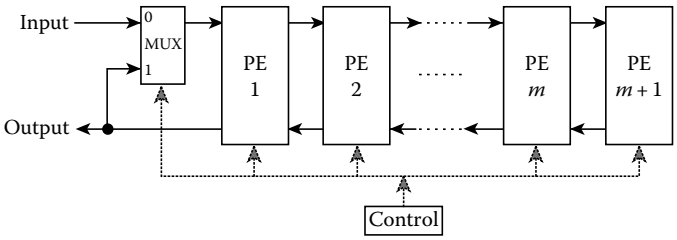


FIGURE 9.105 The overlapped subfilter scheme for IIR filtering.

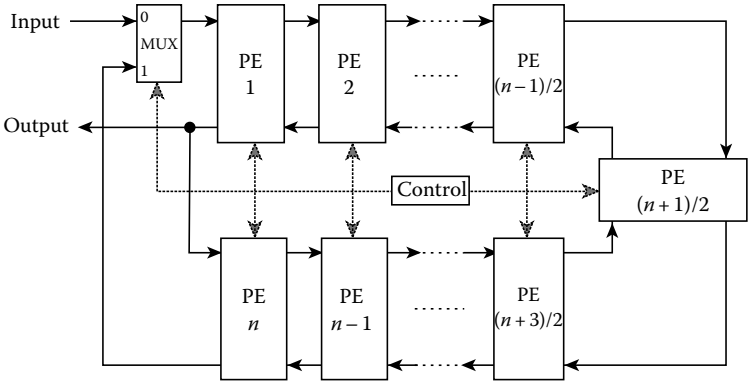


FIGURE 9.106 The systolic ring scheme for IIR filtering.

9.5.3 Systolic Word and Bit-Level Designs

Previous discussions on systolic array designs have taken place at the word level. This is to say that the smallest data or control item exchanged between pairs of processors is constituted by a ''word'' representable by B bits. Each processor in a word-level system has the capability of performing word-level operations. Some may be as complex as floating point multiplications or others as simple as square root operations, etc. The systolic array approach can be applied at various different levels beyond the word level, according to what is sometimes referred to as ''granularity'' of the algorithm description. Systolic arrays and associated dependence graphs can, in fact, be defined at high levels of description, in which each individual processor can, in principle, be a whole mainframe computer or even a separate parallel processor array. The communication between processors thus takes the form of complex protocols and entire data sequences. According to the same principle, the algorithm description can also be done at the lowest level of operation, namely, at the bit level, at which each processor is a simple latched logic gate, capable of performing a logic binary operation. The exchanged data and control also take the form of binary digits. The different approaches due to the different granularity of description have different merits and can be advantageously used in various circumstances or at different steps of the design.

These considerations bring to one possible design strategy, namely, the ''hierarchical systolic design.'' In this approach the complete systolic design is broken down to a sequence of hierarchical steps, each of which define the algorithm at different levels of granularity. At first, the higher level of description is adopted, the relative dependence graph is drawn, and, after suitable projection and scheduling, a high-level systolic architecture is defined. Subsequently, each high-level processor is described in terms of finer scale operations. Dependence graph and systolic architecture corresponding to these operations are produced and embedded in the higher level structure previously obtained. The process can continue down to the desired level of granularity. The simplest form of hierarchical design implies two steps. The first step involves the design of the word-level architecture. Second, the operations performed by each

work-level processor are described at bit level. The corresponding bit-level arrays are then nested into the word-level array, after ensuring that data flows at both levels are fully compatible.

The hierarchical approach has the merits of reducing the complexity of each step of the design. The dependence graphs involved usually have reduced dimensionality (thus, are more manageable), and the procedure is essentially recursive. The drawback of a hierarchical design is that it implicitly introduces somewhat arbitrary boundaries between operations, thereby reducing the set of resulting architectures. An approach that leaves all options open is to consider the algorithm at bit level from the outset. This approach has led to new insights and novel architectures. The price to pay is that the designer must deal with dependence graphs of higher dimensionality. As an example, the dependence graph of the inner product between two N-vectors, $c = \Sigma_{i=0}^{N-1} a_i b_i$ is two dimensional. If the same inner product is written at bit level, i.e., $c_k = \Sigma_{i=0}^{N-1} \Sigma_{j=0}^{B-1} a_{i,j} b_{i,k-j} +$ carries, $k = 0, \ldots, B-1$, then it produces a three-dimensional dependence graph.

Examples of the two design procedures applied to the convolution problem are considered below. First, consider the factors that can make bit-level design advantageous:

- *Regularity*. Most bit-level arrays are highly regular. Only relatively simple cells need to be designed and tested. The communication pattern is simple and regular. Neighbor-to-neighbor connections allow high packing density and low transmission delays.
- *High pipeline rate*. Because the individual cells have reduced computation time (on the order of the propagation delay through a few gates), the overall throughput can be made very high.
- *Inexpensive fault tolerance*. The use of bypass circuitry can be made without wasting too much of the silicon area.

It must be borne in mind that bit-level arrays realistically cannot be operated in wavefront array mode because the interprocessor hand-shaking protocols would be too expensive as compared to the data exchange. A good clock signal distribution is therefore needed to synchronize the array operations. In systolic arrays, unlike synchronous architectures of a different sort, only the incremental clock skew must be minimized by suitably designing the clock signal distribution lines. This problem may become particularly delicate in bit-level arrays, where the number of processors involved is very high.

9.5.3.1 Bit-Level Design of a Serial Convolver

Bit-level systolic design was first proposed by McCanny and McWhirter [29]. Subsequently, they and others have applied this technique to various algorithms. As a simple example, consider the bit-level design of a serial convolver. The word-level output of an N point convolver can be written as $y_k = \Sigma_{i=0}^{N-1} a_i x_{k-i}$, $k = 0, 1, \ldots$, where $\{a_i\}_{i=0}^{N-1}$ is a given set of coefficients and x_i, $i = 0, 1, \ldots$, is a sequence of input data. Coefficients and data values are assumed to be B-bit words. The word-level dependence graph is shown in Figure 9.107, together with one possible systolic realization. In this case, the coefficients are permanently stored in each individual cell. I/O values are propagated in opposite directions. In each PE, the corresponding coefficient is multiplied by the incoming data value. This product is added to the partial output value and the accumulated result is propagated forward. Each cell performs the simple multiply and add operation expressed by $y_{k,i+1} \leftarrow y_{k,i} + a_i x_{k-i}$, $y_k = y_{k,N}$.

According to the hierarchical approach, one must now proceed to determine the dependence graph corresponding to the bit-level description of the multiply-and-add operation. The complete dependence graph can be subsequently obtained by embedding the finer scale graph into the higher level graph. If both a_i and x_i are B-bit binary numbers, then the jth bit of y_{ki}, can be computed according to $y_{k,i,j} = y_{k,i,j} + s_{i,k,j}, s_{i,k,j} + \Sigma_{l=0}^{B-1} a_{i,l} x_{k-i,j-l} +$ carries, where $a_{i,l}$ and $x_{i,l}$, $l = 0, \ldots, B-1$, represent the lth bit of a_i and x_i. The dependence graph corresponding to this operation is given in Figure 9.108, where subscripts only indicate the bit position, and $B = 3$. Note that this graph is quite similar to the graph corresponding to a convolver, apart from the carry bits, which are taken care of by the insertion of an additional row of cells.

The combined dependence graph, obtained from the word dependence graph of Figure 9.107, in which each cell is replaced by the bit-level dependence graph of Figure 9.108, is given in Figure 9.109.

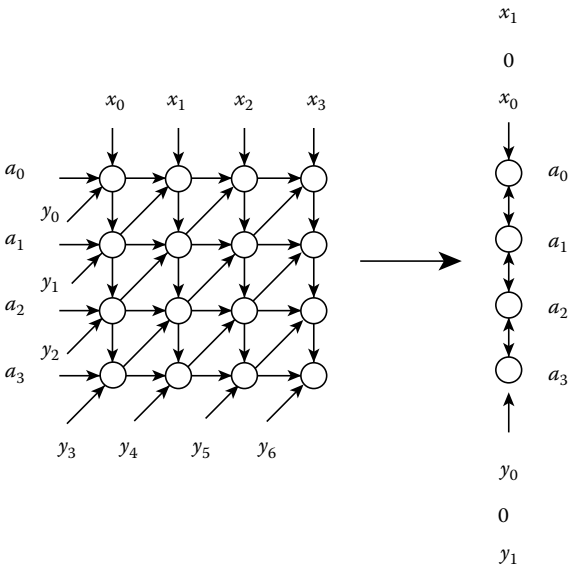


FIGURE 9.107 World-level dependence graph of an N point convolution operation with one possible systolic realization.

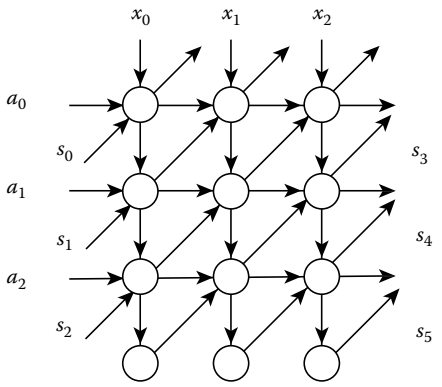


FIGURE 9.108 Bit-level dependence graph corresponding to the multiply-and-add operation.

The data flows are fully compatible at both word and bit levels. At this point, a full two-dimensional bit-level systolic array can be obtained from the final dependence graph by simply replacing each node with latched full adder cells. Different linear systolic implementations can be obtained by projecting the combined dependence graph along various directions. One possibility is again to keep the coefficients residents in individual cells, and have input data bits and accumulated results propagate in opposite directions. The schematic representation of the systolic array with these features is drawn in Figure 9.109. Judgment about the merits of different projections involves desired data movement, I–O considerations, throughput rate, latency time, efficiency factor (ratio of idle time to busy time per cell), etc.

As discussed previously, the convolution operation can be described at bit level from the very beginning. In this case the expression for the jth bit of the kth output can be expressed as follows:

$$y_{k,j} = \sum_{i=0}^{N-1} \sum_{l=0}^{B-1} a_{i,l} x_{k-i,j-1} + \text{carries} \tag{9.23}$$

By using this expression as a starting point, one is capable of generating a number of feasible systolic realizations potentially much larger than what is attainable from the two-step hierarchical approach. The reason for this can be simply understood by nothing that in this formulation no arbitrary precedence relationship is imposed between the two summations on i and l, whereas earlier we required that the summation on l would always ''precede'' the summation on i. The result is a fairly complicated three-dimensional dependence graph of size $N \times B \times$ number of inputs, as shown in Figure 9.110. Observe that

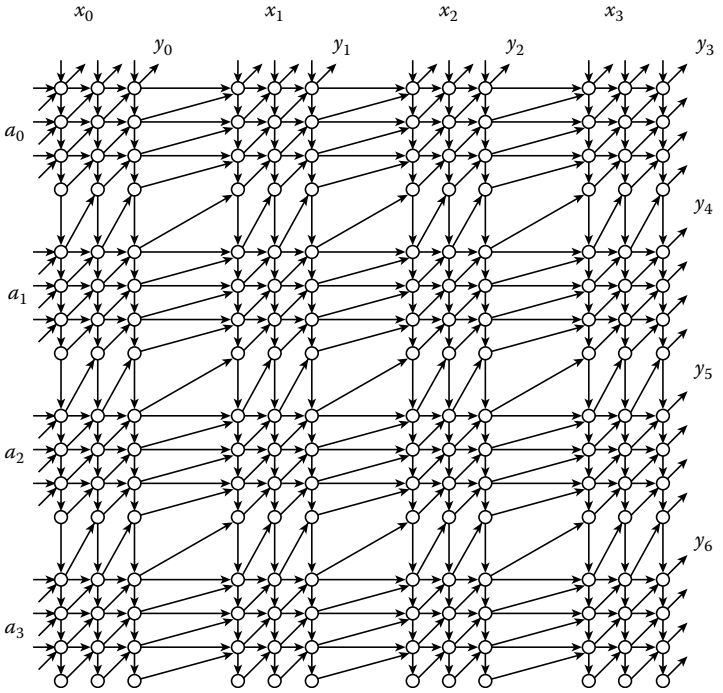


FIGURE 9.109 Bit-level dependence graph for convolution obtained by embedding the bit-level graph into the word-level graph.

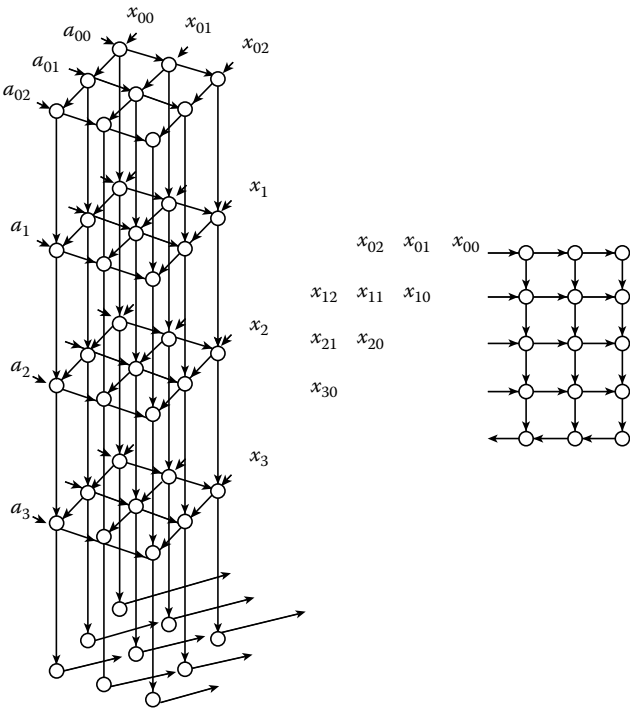


FIGURE 9.110 General three-dimensional bit-level dependence graph for convolution.

the bottom level of the dependence graph corresponds to the summation over l in Equation 9.23. In the same figure a schematic two-dimensional bit-level systolic realization of the algorithm is given, in which the coefficient bits are held in place. Projections along different directions have different characteristics and may be considered preferable in different situations. The choice ultimately must be made according to given design constraints or to efficiency requirements.

The concept of bit-level design, as considered here, can be applied to a large variety of algorithms. Indeed, it has generated a number of architectures, including FIR/IIR filters, arrays for inner product computation, median filtering, image processing, eigenvalue problems, Viterbi decoding, etc.

9.5.4 Recursive LSs Estimation

9.5.4.1 LSs Estimation

The LS technique constitutes one of the most basic components of all modern signal processing algorithms dealing with linear algebraic and optimization of deterministic and random signals and systems. Specifically, some of the most computationally intensive parts of modern spectral analysis, beam formation, direction finding, adaptive array, image restoration, robotics, data compression, parameter estimation, and KF all depend crucially on LS processing.

Regardless of specific application, an LS estimation problem can be formulated as $Ax \approx y$, where the $m \times n$ data matrix A and the $m \times 1$ data vector y are known, and we seek the $n \times 1$ desired solution x. In certain signal processing problems, rows of A are composed of sequential blocks of lengths n taken from a one-dimensional sequence of observed data. In other n-sensor multichannel estimation problems, each column of A denotes the sequential outputs of a given sensor. In all cases, the desired solution x provides the weights on the linear combinations of the columns of A to optimally approximate the observed vector y in the LS sense. When $m = n$ and A is nonsingular, then an exact solution for x exists. The Gaussian elimination method provides an efficient approach for determining this exact solution. However, for most signal processing problems, such as when there are more observations than sensors, and thus $m > n$, then no exact solution exists. The optimum LS solution $\hat{x}$ is defined by $\|A\hat{x} - y\| = \min_x \|Ax - y\|$. The classical approach in LS solution is given by $\hat{x} = A^+ y$, where A^+ is the pseudo-inverse of A defined by $A^+ = (A^{\mathrm{T}} A)^{-1} A^{\mathrm{T}}$. The classical LS approach is not desirable from the complexity, finite precision sensitivity, and processing architecture points of views. This is due to the need for a matrix inversion, the increase of numerical instability from "squaring of the condition number" in performing the $A^{\mathrm{T}}A$ operation, and the block nature of the operation in preventing a systolic update processing and architecture for real-time applications.

The QRD approach provides a numerically stable technique for LS solution that avoids the objections associated with the classical approach. Consider a real-valued $m \times n$ matrix A with $m \geq n$ and all the columns are linearly independent (i.e., rank $A = n$). Then, from the QRD, we can find a $m \times m$ orthogonal matrix Q such that $QA = \overline{R}$. The $m \times n$ matrix $\overline{R} = [R^{\mathrm{T}}, 0^{\mathrm{T}}]^{\mathrm{T}}$ is such that R is an $n \times n$ upper triangular matrix (with nonzero diagonal elements) and 0 is an all-zero $(m - n) \times n$ matrix. This upper triangularity of R is used crucially in the following LS solution problem.

Because the l_2 norm of any vector is invariant with respect to an orthogonal transformation, an application of the QRD to the LS problem yields $\|Ax - y\|^2 = \|Q(Ax - y)\|^2 = \|\overline{R}x - f\|^2$, where f is an $m \times 1$ matrix given by $f = Qy = [u^{\mathrm{T}}, v^{\mathrm{T}}]^{\mathrm{T}}$. Denote $e = Ax - y$ as the "residual" of the LS problem. Then, the previous LS problem is equivalent to $\|e\|^2 = \|Ax - y\|^2 = \|[Rx, 0x]^{\mathrm{T}} - [u^{\mathrm{T}} - v^{\mathrm{T}}]^{\mathrm{T}}\|^2 = \|Rx - u\|^2 + \|v\|^2$. Because R is a nonsingular upper triangular square matrix, the back substitution procedure of the Gaussian elimination method can be used to solve for the exact solution $\hat{x}$ of $R\hat{x} = u$. Finally, the LS problem reduces to $\min_x \|Ax - y\|^2 = \|A\hat{x} - y\|^2 = \|-R\hat{x} - f\|^2 = \|R\hat{x} - u\|^2, + \|v\|^2 = \|v\|^2$. For the LS problem, any QRD technique such as the Gram–Schmidt method, the modified-Gram–Schmidt (MGS) method, the Givens transformation, and the Householder transformation is equally valid for

finding the matrix R and the vector v. For a systolic implementation, the Givens transformation yields the simplest architecture, but the MGS and Householder transformation techniques are also possible with slight advantages under certain finite precision conditions.

9.5.4.2 Recursive LSs Estimation

The complexity involved in the computation of the optimum residual $\hat{e}$ and the optimum LS solution vector $\hat{x}$ can become arbitrarily large as the number of samples in the column vectors of A and y increases. In practice, we must limit m to some finite number greater than the number of columns n. Two general approaches in addressing this problem are available. In the ''sliding window'' approach, we periodically incorporate the latest observed set of data (i.e., ''updating'') and possibly remove an older set of data (i.e., ''downdating''). In the ''forgetting factor'' approach, a fixed scaling constant with a magnitude between 0 and 1 is multiplied against the R matrix and thus exponentially forget older data. In either approach, we find the optimum LS solution weight vector $\hat{x}$ in a recursive LSs manner. As the statistics of the signal change over each window, these $\hat{x}$ vectors change ''adaptively'' with time. This observation motivates the development of a recursive LSs solution implemented via the QRD approach. For simplicity, we consider only the updating aspects of the sliding window recursive LSs problem.

Let m denote the present time of the sliding window of size m. Consider the $m \times n$ matrix $A(m)$, the $m \times 1$ column vector $y(m)$, the $n \times 1$ solution weight column vector $x(m)$, and the $m \times 1$ residual column vector $e(m)$ expressed in terms of their values at time $m-1$ as $A(m) = [\alpha\ (1), \ldots, \alpha\ (m)]^{\mathrm{T}} = [A(m-1)^{\mathrm{T}}, \alpha\ (m)]^{\mathrm{T}}$, $y(m) = [y_1, \ldots, y_m]^{\mathrm{T}}\ [y(m-1)^{\mathrm{T}}, y^{\mathrm{T}}m]^{\mathrm{T}}$, $x(m) = [x_1(m), \ldots, x_n(m)]^{\mathrm{T}}$, and $e(m)\ A(m) \times (m) - y(m) = [e_1(m), \ldots, e_n(m)]^{\mathrm{T}}$. By applying the orthogonal matrix $Q(m) = [Q_1(m)^{\mathrm{T}}, Q_2(m)^{\mathrm{T}}]^{\mathrm{T}}$ of the QRD of the $m \times n$ matrix $A(m)$, we obtain $Q(m)\ A(m) = [R(m)^{\mathrm{T}},\ 0^{\mathrm{T}}]^{\mathrm{T}} = R_0(m)$ and $Q(m)y(m) = [Q_1(m)^{\mathrm{T}}, Q_2(m)^{\mathrm{T}}]^{\mathrm{T}}\ y(m) = [u(m)^{\mathrm{T}},\ v(m)^{\mathrm{T}}]^{\mathrm{T}}$. The square of the l_2 norm of the residual e is then given by $\in(m) = \|e(m)\|^2 = \|A(m)x(m) - y(m)\|^2 = \|Q(m)(A(m)x(m) - y(m))\|^2 = \|R(m)x(m) - u(m)\|^2 + \|v(m)\|^2$. The residual is minimized by using the back substitution method to find the optimum LS solution $\hat{x}(m)$ satisfying $R(m)\hat{x}(m) = u(m) = [u_1(m), \ldots, u_n(m)]^{\mathrm{T}}$. It is clear that the optimum residual $\hat{e}(m)$ is available after the optimum LS solution $\hat{x}(m)$ is available as seen from $\hat{e}(m) = A(m)\hat{x}(m) - y(m)$. It is interesting to note that it is not necessary to first obtain $\hat{x}(m)$ explicitly and then solve for $\hat{e}(m)$ as shown earlier. It is possible to use a property of the orthogonal matrix $Q(m)$ in the QRD of A and the vector $y(m)$, to obtain $\hat{e}(m)$ explicitly. Specifically, note $\hat{e}(m) = A(m)\hat{x}(m) - y(m) = Q_1(m)^{\mathrm{T}}\ R(m)\hat{x}(m) - y(m) = [Q_1(m)^{\mathrm{T}}\ Q_1(m) - I_m]y(m) = -Q_2(m)^{\mathrm{T}}\ Q_2(m)y(m) = -Q_2(m)^{\mathrm{T}}v(m)$. This property is used explicitly in the following systolic solution of the last component of the optimum residual.

9.5.4.3 Recursive QRD

Consider the recursive solution of the QRD. First, assume the decomposition at step $m-1$ has been completed as given by $Q(m-1)A(m-1) = [R(m-1)^{\mathrm{T}}, 0^{\mathrm{T}}]^{\mathrm{T}}$ by using a $(m-1) \times (m-1)$ orthogonal matrix. Next, define a new $m \times m$ orthogonal transformation $T(m) = [Q(m-1), 0; 0, 1]$. By applying $T(m)$ on the new $m \times n$ data $A(m)$, which consists of the previously available $A(m-1)$ and the newly available row vector $\alpha(m)^{\mathrm{T}}$ we have

$$T(m)A(m) = \begin{bmatrix} Q(m-1) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A(m-1) \\ \alpha(m)^T \end{bmatrix} = \begin{bmatrix} Q(m-1)A(m-1) \\ \alpha(m)^T \end{bmatrix}$$
$$= \begin{bmatrix} R(m-1) \\ 0 \\ \alpha(m)^T \end{bmatrix} = R_1(m)$$

While $R(m-1)$ is an $n \times n$ upper triangular matrix, $R_1(m)$ does not have the same form as the desired $R_0(m) = [R(m)^{\mathrm{T}}, 0^{\mathrm{T}}]^{\mathrm{T}}$ where $R(m)$ is upper triangular.

9.5.4.4 Givens Orthogonal Transformation

Next, we want to transform $R_1(m)$ to the correct $R_0(m)$ form by an orthogonal transformation $G(m)$. While any orthogonal transformation is possible, we will use the Givens transformation approach due to its simplistic systolic array implementation. Specifically, denote $G(m) = G_n(m)G_{n-1}(m)\ldots G_1(m)$, where $G(m)$ as well as each $G_i(m)$, $i = 1,\ldots n$, are all $m \times m$ orthogonal matrices. Define

$$G_i(m) = \begin{array}{c} \\ 1 \\ \\ i \\ \\ m \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & & i & & m \end{array} \\ \begin{bmatrix} 1 & & 0 & & 0 \\ & 1 & & & \\ 0 & & c_i(m) & & s_i(m) \\ & & & 1 & \\ 0 & & -s_i(m) & & c_i(m) \end{bmatrix} \end{array}, \quad i = 1,\ldots,n$$

as a $m \times m$ identity matrix, except that the (i, i) and (m, m) elements are specified as $c_i(m) = \cos\theta_i(m)$, where $\theta_i(m)$ represents the rotation angle at the ith iteration, the (i, m) element as $s_i(m) = \sin\theta_i(m)$, and the (m, i) element as $-S_i(m)$. By cascading all the $G_i(m)$, $G(m)$ can be reexpressed as

$$G(m) = \begin{bmatrix} k(m) & 0 & d(m) \\ 0 & I_{m-n-1} & 0 \\ h^T(m) & 0 & \gamma(m) \end{bmatrix}$$

where $k(m)$ is $n \times n$, $d(m)$ and $h(m)$ are $n \times 1$, and $\gamma(m)$ is 1×1. In general $k(m)$, $d(m)$, and $h(m)$ are quite involved functions of $c_i(m)$ and $s_i(m)$, but $\gamma(m)$ is given simply as $\gamma(m) = \prod_{i=1}^{n} c_i(m)$ and will be used in the evaluation of the optimum residual.

Use $G(m)$ to obtain $G(m)T(m)A(m) = G(m)R_1(m)$. In order to show the desired property of the n orthogonal transformation operations of $G(m)$, first consider

$$G_1(m)R_1(m) = \begin{bmatrix} c_1(m) & & & & s_1(m) \\ & 1 & & & \\ & & 1 & & \\ & & & 1 & \\ s_1(m) & & & & c_1(m) \end{bmatrix} \begin{bmatrix} x & x & \cdots & x \\ 0 & x & \cdots & x \\ & & \cdots & x \\ 0 & 0 & \cdots & x \\ x & x & \cdots & x \end{bmatrix}$$

$$= \begin{bmatrix} x & x & \cdots & x \\ 0 & x & \cdots & x \\ & & \cdots & x \\ 0 & 0 & \cdots & x \\ 0 & x & \cdots & x \end{bmatrix}$$

In the preceding expression, an x denotes some nonzero valued element. The purpose of $G_1(m)$ operating on $R_1(m)$ is to obtain a zero at the $(m, 1)$ position without changing the $(m-2) \times n$ submatrix from the second to the $(m-1)$st rows of the r.h.s. of the expression. In general, at the ith iteration, we have

$$G_i(m) \begin{bmatrix} x & x & . & . & . & x \\ & x & . & . & . & x \\ & & & & & x \\ & & & \vdots & & \\ 0 & 0 & . & . & . & 0 \\ 0 & 0 & 0 & x & x & x \end{bmatrix} = \begin{bmatrix} x & x & . & . & . & x \\ & x & . & . & . & x \\ & & & & & x \\ & & & \vdots & & \\ 0 & 0 & . & . & . & 0 \\ 0 & 0 & 0 & 0 & x & x \end{bmatrix}$$

$i-1$ $\qquad$ i

The preceding zeroing operation can be explained by noting that the Givens matrix $G_i(m)$ operates as a $(m-2)\times(m-2)$ identity matrix on all the rows on the right of it except the ith and the mth rows. The crucial operations at the ith iteration on these two rows can be represented as

$$\begin{bmatrix} c & s \\ -s & c \end{bmatrix} \quad \begin{bmatrix} 0 & \cdots & 0 & r_i & r_{i+1} & \cdots & r_n \\ 0 & \cdots & 0 & \underset{i}{a_i} & a_{i+1} & \cdots & a_n \end{bmatrix}$$

$$= \begin{bmatrix} 0 & \cdots & 0 & r_i^T & r_{i+1}^T & \cdots & r_n^T \\ 0 & \cdots & 0 & \underset{i}{0} & a_{i+1}^T & \cdots & a_n^T \end{bmatrix}$$

For simplicity of notation, we suppress the dependencies of i and m on c and s. Specifically, we want to force $a_i^{\mathrm{T}}=0$ as given by $0=a_i^{\mathrm{T}}=-sr_i+ca_i$. In conjunction with $c^2+s^2=1$, this requires $c^2=r_i^2/(a_i^2+r_i^2)$ and $s^2=a_i^2/(a_i^2/r_i^2)$. Then $r_i^{\mathrm{T}}=cr_i+sa_i=\sqrt{(a_i^2+r_i^2)}$, $c=r_i/r_i^{\mathrm{T}}$, and $s=a_i/r_i^{\mathrm{T}}$. This shows from the individual results of $G_1(m)$, $G_2(m),\ldots,$ $G_n(m)$, the overall results yield $Q(m)\ A(m)=G(m)\overline{R}(m)=[R(m)^{\mathrm{T}},\ 0^{\mathrm{T}}]^{\mathrm{T}}=R_0(m)$, with $Q(m)=G(m)T(m)$.

9.5.4.5 Recursive Optimal Residual and LS Solutions

Consider the recursive solution of the last component of the optimum residual $\hat{e}(m)=[\hat{e}_1(m),\ldots,\hat{e}_m(m)]^{\mathrm{T}}=-Q_2(m)^{\mathrm{T}}\,v(m)=-Q_2(m)^{\mathrm{T}}\,[v_1(m),\ldots,v_m(m)^{\mathrm{T}}]$. Because $Q_2(m)=[Q_2(m-1),\ 0;\ h(m)^{\mathrm{T}}\ Q_1(m-1),\ \gamma(m)]$, then $\hat{e}(m)=[\hat{e}_1(m),\ldots,\hat{e}_m(m)]^{\mathrm{T}}=Q_2(m)=-[Q_2^{\mathrm{T}}(m-1),\ Q_1^{\mathrm{T}}(m-1)h(m);\ 0,\ \gamma(m)]$ $[v_1(m),\ldots,v_m(m)]$. Thus, the last component of the optimum residual is given by $\hat{e}(m)=-\gamma(m)$ $v_m(m)=-\prod_{i=1}^{n}c_i(m)v_m(m)$, which depends on all the products of the cosine parameters $c_i(m)$ in the Givens QR transformation, and $v_m(m)$ is just the last component of $v(m)$, which is the result of $Q(m)$ operating on $y(m)$.

As considered earlier, the LS solution $\hat{x}$ satisfies the triangular system of equations. After the QR operation on the extended matrix $[A(m),\ Y(m)]$, all the $r_{ij}, j\geq i=1,\ldots,n$ and $u_i, i=1,\ldots,n$ are available. Thus, $\{\hat{x}_1,\ldots,\hat{x}_n\}$ can be obtained by using the back substitution method of $\hat{x}_i=(u_i-\Sigma_{j-i+1}^{n}r_{ij}\hat{x}_j/r_{ij})$, $i=n,\ n-1,\ldots,\ 1$. Specifically, if $n=1$, then $\hat{x}_1=u_1/r_{11}$. If $n=2$, then $\hat{x}_2=u_2/r_{22}$ and $\hat{x}_1=u_1-r_{12}\hat{x}_2/r_{11}=u_1/r_{11}-u_2r_{12}/r_{11}r_{22}$. If $n=3$, then $\hat{x}_3=u_3/r_{33}, \hat{x}_2=u_2-r_{23}x_3/r_{22}=u_2/r_{22}-r_{23}u_3/r_{22}r_{23}$, and $\hat{x}_1=u_1-r_{12}\hat{x}_2-r_{13}\hat{x}_3/r_{11}=u_1/r_{11}-r_{12}u_2/r_{11}r_{22}+u_3[-r_{13}/r_{11}r_{33}+r_{12}r_{23}/r_{11}r_{22}r_{33}]$.

9.5.4.6 Systolic Array Implementation for QRD and LS Solution

The recursive QRD considered above can be implemented on a two-dimensional triangular systolic array based on the usage of four kinds of processing cells. Figure 9.111a shows the boundary cell for the generation of the sine and cosine parameters, s and c, needed in the Givens rotations. Figure 9.111b shows the internal cell for the proper updating of the QRD transformations. Figure 9.111c shows the single output cell needed in the generation of the last component of the optimal residual $\hat{e}_m(m)$ as well as the optimal LS solution $\hat{x}(m)$. Figure 9.111d shows the delay cell which performs a unit time delay for proper time skewing in the systolic processing of the data.

Figure 9.112 shows a triangular systolic array capable of performing the recursive QRD for the optimal recursive residual estimation and the recursive LSs solution by utilizing the basic processing cells in Figure 9.111. In particular, the associated LS problem uses an augmented matrix $[A,y]$ consisting of the $m\times n$ observed data matrix A and the $m\times 1$ observed vector y. The number or processing cells in the triangular array consists of n boundary cells, $n(n+1)/2$ internal cells, one output cell, and n delay cells.

The input to the array in Figure 9.112 uses the augmented matrix

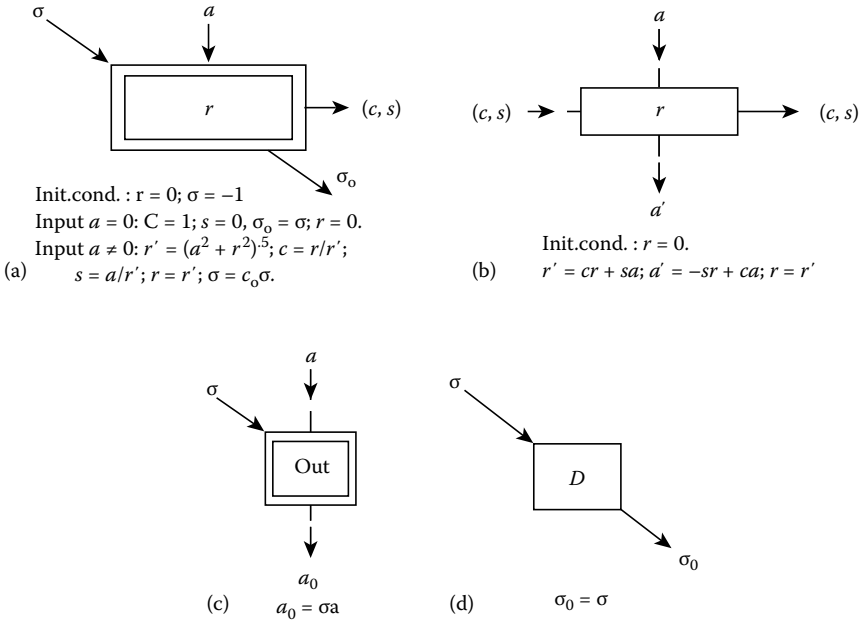


FIGURE 9.111 (a) Boundary cell; (b) internal cell; (c) output cell; and (d) delay cell.

$$[A,Y] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & y_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & y_2 \\ & & \vdots & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} & y_m \end{bmatrix}$$

skewed in a manner such that each successive column from left to right is delayed by a unit time as given by

$$\begin{bmatrix} a_{11} & 0 & & \cdots & 0 & 0 \\ a_{21} & a_{12} & 0 & \cdots & 0 & 0 \\ a_{31} & a_{22} & a_{13} & \cdots & 0 & 0 \\ & & & \vdots & & \\ a_{n1} & a_{(n-1)2} & & & a_{1n} & 0 \\ a_{(n+1)1} & a_{n2} & & & a_{2n} & y_1 \\ & & & \vdots & & \\ a_{m1} & a_{(m-1)2} & & & a_{(m-n+1)n} & y_{m-n} \\ & & & \vdots & & \\ 0 & & & & a_{mn} & y_{m-1} \\ 0 & & & & 0 & y_m \end{bmatrix} \begin{matrix} k=1 \\ 2 \\ 3 \\ \vdots \\ n \\ n+1 \\ \vdots \\ m \\ \vdots \\ m+n-1 \\ m+n \end{matrix}$$

We see that at time k, input data consists of the kth row of the matrix, and moves down with increasing time. However, in Figure 9.112, purely for drawing purpose in relation to the position of the array, the relevant rows of data are drawn as moving up with increasing k.

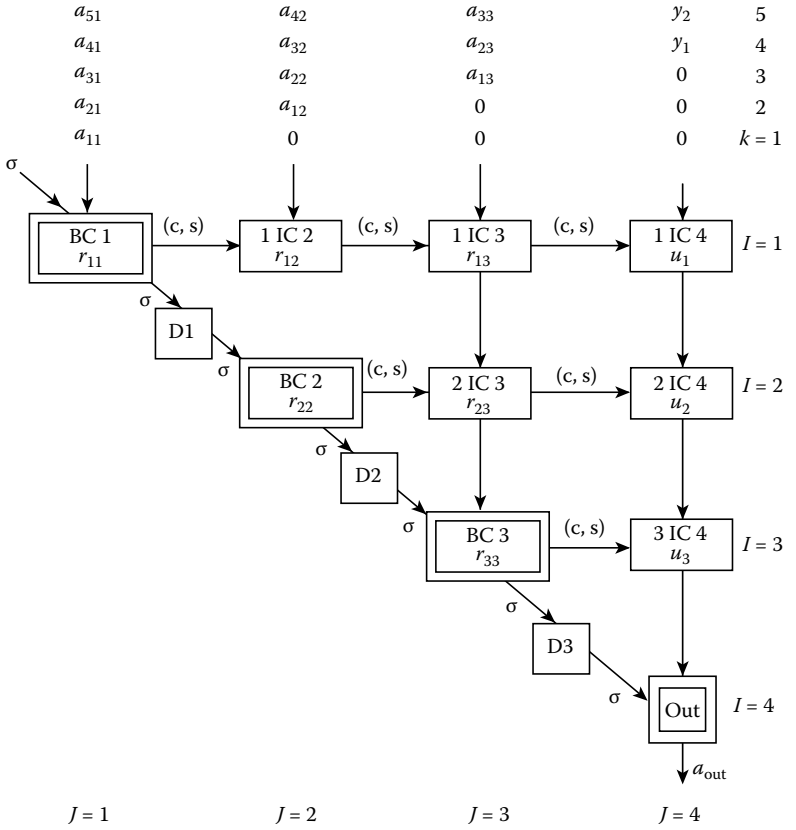


FIGURE 9.112 Triangular systolic array implementation of an $n = 3$, QRD-recursive, LSs solver.

Consider some of the iterative operations of the QRD for the augmented matrix $[A, y]$ for the systolic array in Figure 9.112. At time $k = 1$, a_{11} enters BC 1 and results in $c = 0$, $s = 1$, and $r_{11} = a_{11}$. All other cells are inactive. At $k = 2$, a_{21} enters BC 1, with the results $c = a_{11}\sqrt{(a_{11}^2 + a_{21}^2)}$, $s = a_{21}\sqrt{(a_{11}^2 + a_{21}^2)}$,$r_{11} = \sqrt{a_{11}^2 + a_{21}^2}$. This r_{11} corresponds to that of r_i^{T}, while the preceding c and s correspond to the c and s in the Givens transformation. Indeed, the new a_i^{T} is zero and does not need to be saved in the array. Still, at $k = 2$, a_{12} enters 1 IC 2 and outputs $a^{\mathrm{T}} = 0$ and $r_{12} = a_{12}$. At $k = 3$, a_{13} enters BC 1, and the Givens rotation operation continues where the new r_i is given by the previously processed r_i^T and a_i is now given by a_{13}. Meanwhile, a_{22} enters at 1 IC 2. It outputs $a^{\mathrm{T}} = -a_{21}a_{12}/\sqrt{(a_{11}^2 + a_{21}^2)} + -a_{22}a_{21}/\sqrt{(a_{11}^2 + a_{21}^2)}$, which corresponds to that of a_{i+1}^T, and $r_{12} = a_{11}a_{12}/\sqrt{(a_{11}^2 + a_{21}^2)} + -a_{21}a_{22}/\sqrt{(a_{11}^2 + a_{21}^2)}$, which corresponds to that of r_{i+1}^T. In general, the top (i.e., $I = 1$) row of the processing cells performs Givens rotation by using the first row to operate on the second, third,..., mth rows (each row with $n + 1$ elements), such that the $\{21, 31, \ldots, m1\}$ locations in the augmented matrix are all zeroed. The next row ($I = 2$) of cells uses the second row to operate on the third,..., mth rows (each row with n elements), such that locations at $\{32, 42, \ldots, m2\}$ are zeroed. Finally, at row $I = n$, by using the nth row to operate on the $(n + 1)$st,..., mth rows, elements at locations $\{(n + 1)n, (n + 2)n, \ldots, mn\}$ are zeroed. We also note that the desired cosine values in $\gamma(m)$ are being accumulated by c along the diagonal of the array. Delay cells $\{D_1, D_2, \ldots, D_n\}$ are used to provide the proper timing along the diagonal.

The cell BC 1 (at $I = J = 1$) terminates in the QR operation at time $k = m$, while the cell at $I = 1$ and $J = 2$ terminates at $k = m + 1$. In general, the processing cell at location (I, J) terminates at

$k = I + J + m - 2$. In particular, the last operation in the QRD on the augmented matrix is performed by the cell at $I = n$ and $J = n + 1$ at time $k = 2n + m - 1$. Then, the last component of the optimum residual $e_m(m)$ exits the output cell at time $k = 2n + m$.

After the completion of the QRD obtains the upper triangular system of equation, we can "freeze" the r_{IJ} values in the array to solve for the optimum LS solution $\hat{x}$ by the back substitution method. Specifically, we can append $[I_n, 0]$, where I_n is a $n \times n$ identity matrix and 0 is an $n \times 1$ vector of all zeroes, to the bottom of the augmented matrix $[A, y]$. Of course, this matrix is skewed as before when used as input to the array. In particular, immediately after the completion of the QR operation at BC 1, we can input the unit value at time $k = m + 1$. This is stage 1 of the back substitution method. Due to skewing, a unit value appears at the $I = 1$ and $J = 2$ cell at stage 3. Finally, at stage $(2n - 1)$, which is time $k = m + 2n - 1$, the last unit value appears at the $I = n$ and $J = 1$ cell. For our example of $n = 3$, this happens at stage 5. The desired LS solution $\hat{x}_1$ appears at stage $(2n + 1)$ (i.e., stage 7 for $n = 3$), which is time $k = 2n + m + 1$, while the last solution $\hat{x}_n$ appears at stage $3n$ (i.e., stage 9 for $n = 3$), which is time $k = 3n + m$. The values of$\{\hat{x}_1, \hat{x}_2, \hat{x}_3\}$ at the output of the systolic array are identical to those given by the back substitution method solution of the LS problem.

9.5.5 Kalman Filtering

KF was developed in the late 1950s as a natural extension of the classical Wiener filtering. It has profound influence on the theoretical and practical aspects of estimation and filtering. It is used almost universally for tracking and guidance of aircraft, satellites, GPS, and missiles as well as many system estimation and identification problems. KF is not one unique method, but is a generic name for a class of state estimators based on noisy measurements. KF can be implemented as a specific algorithm on a general-purpose mainframe/mini/microcomputer operating in a batch mode, or it can be implemented on dedicated system using either DSP, ASIC, or custom VLSI processors in a real-time operating mode.

Classically, an analog or a digital filter is often viewed in the frequency domain having some low-pass, bandpass, high-pass, etc. properties. A KF is different from the classical filter in that it may have multiple inputs and multiple outputs with possibly nonstationary and time-varying characteristics performing optimum states estimation based on the unbiased minimum variance estimation criterion.

In the following discussions, we first introduce the basic concepts of KF, followed by various algorithmic variations of KF. Each version has different algorithmic and hardware complexity and implementational implications. Because there are myriad of KF variations, we then consider two simple systolic versions of KF.

9.5.5.1 Basic KF

The KF model consists of a discrete-time linear dynamical system equation and a measurement equation. A linear discrete-time dynamical system with $n \times 1$ state vector $x(k + 1)$, at time $k + 1$, is given by $x(k + 1) = \mathrm{A}(k)x(k) + B(k)u(k) + w(k)$, where $x(k)$ is the $n \times 1$ state vector at time k, $A(k)$ is an $n \times n$ system coefficient matrix, $B(k)$ is an $n \times p$ control matrix, $u(k)$ is a $p \times 1$ deterministic vector, which for some problems may be zero for all k, and $w(k)$ is an $n \times 1$ zero-mean system noise vector with a covariance matrix $W(k)$. The input to the KF is the $m \times 1$ measurement (also called observation) vector $y(k)$, modeled by $y(k) = C(k)x(k) + v(k)$, where $C(k)$ is an $m \times n$ measurement coefficient matrix, and $v(k)$ is a $m \times 1$ zero-mean measurement noise vector with an $m \times m$ positive-definite covariance matrix $V(k)$. The requirement of the positive-definite condition on $V(k)$ is to guarantee the Cholesky (square root) factorization of $V(k)$ for certain KF algorithms. In general, we will have $m \leq n$ (i.e., the measurement vector dimension is less than or equal to that of the state vector dimension). It is also assumed that $w(k)$ is uncorrelated to $v(k)$. That is, $E\{w(i)v(j)^{\mathrm{T}}\} = 0$. We also assume each noise sequence is white in the sense $E\{w(i)w(j)^{\mathrm{T}}\} = E\{v(i)v(j)^{\mathrm{T}}\} = 0$, for all $i \neq j$.

The KF provides a recursive linear estimation of $x(k)$ under the minimum variance criterion based on the observation of the measurement $y(k)$. Let $\hat{x}(k)$ denote the optimum filter state estimate of $x(k)$ given

measurements up to and including $y(k)$, while $\hat{x}_+(k)$ denotes the optimum predicted state estimate of $x(k)$ given measurement up to and including $y(k-1)$. Then the $n \times n$ "optimum estimation error covariance matrix" is given by $P(k) = E\{(x(k) - \hat{x}(k))(x(k) - \hat{x}(k))^{\mathrm{T}}\}$, while the "minimum estimation error variance" is given by $J(k) = \text{Trace } P(k) = E\{(x(k) - \hat{x}(k))^{\mathrm{T}} (x(k) - \hat{x}(k))\}$. The $n \times n$ "optimum prediction error covariance matrix" is given by $P_+(k) = E\{(x(k) - x_+(k))(x(k) - x_+(k))^{\mathrm{T}}\}$.

The original KF recursively updates the optimum error covariance and the optimum state estimate vector by using two sets of update equations. Thus, it is often called the "covariance KF." The "time update equations" for $k = 1, 2, \ldots$, are given by $x_+(k) = A(k-1)\ \hat{x}(k-1) + B(k-1)u(k-1)$ and $P_+(k) = A(k-1)\ P(k-1)A^{\mathrm{T}}(k-1) + W(k-1)$. The $n \times n$ "Kalman gain matrix" $K(k)$ is given by $K(k) = P_+(k)C^{\mathrm{T}}(k)[C(k)P_+(k)C^{\mathrm{T}}(k) + V(k)]^{-1}$. The "measurement update equations" are given by $\hat{x}(k) = x_+(k) + K(k)(y(k) - C(k)x_+(k))$ and $P(k) = P_+(k) - K(k)P_+(k)$. The first equation shows the update relationship of $\hat{x}(k)$ to the predicted state estimate $x_+(k)$, for $x(k)$ based on $\{\ldots, y(k-2), y(k-2), y(k-1)\}$, when the latest observed value $y(k)$ is available. The second equation shows the update relationship between $P(k)$ and $P_+(k)$. Both equations depend on the $K(k)$, which depends on the measurement coefficient matrix $C(k)$and the statistical property of the measurement noise, covariance matrix $V(k)$. Furthermore, $K(k)$ involves an $m \times m$ matrix inversion.

9.5.5.2 Other Forms of KF

The basic KF algorithm considered above is called the covariance form of KF because the algorithm propagates the prediction and estimation error covariance matrices $P_+(k)$ and $P(k)$. Many versions of the KF are possible, characterized partially by the nature of the propagation of these matrices. Ideally, under infinite precision computations, no difference in results is observed among different versions of the KF. However, the computational complexity and the systolic implementation of different versions of the KF are certainly different. Under finite precision computations, especially for small numbers of bits under fixed point arithmetics, the differences among different versions can be significant. In the following discussions we may omit the deterministic control vector $u(k)$ because it is usually not needed in many problems. In the following chol (.) qr (.), and triu (.) stand for Cholesky factor, QRD, and triangular factor, respectively.

1. **Information filter**. The inverse of the estimation error covariance matrix $P(k)$ is called the information matrix and is denoted by $PI(k)$. A KF can be obtained by propagating the information matrix and other relevant terms. Specifically, the information filter algorithm is given by time updates for $k = 1, 2, \ldots$, of $L(k) = A^{-\mathrm{T}}(k-1)PI(k-1)A^{-1}(k-1) \times [W^{-1}(k-1) + A^{-\mathrm{T}}(k-1)\ PI(k-1)A^{-\mathrm{T}}(k-1)]^{-1}$, $d_+(k) = (I - L(k))A^{-\mathrm{T}}(k-1)PI(k-1)A^{-1}(k-1)$. The measurements updates are given by $d(k) = d_+(k) + C^{\mathrm{T}}(k)V^{-1}(k)y(k)PI(k) = PI_+(k) + C^{\mathrm{T}}(k)V^{-1}(k)C(k)$.
2. **Square-root covariance filter (SRCF)**. In this form of the KF, we propagate the square root of $P(k)$. In this manner, we need to use a lower dynamic range in the computations and obtain a more stable solution under finite precision computations. We assume all three relevant covariance matrices are positive-definite and have the factorized form of $P(k) = S^{\mathrm{T}}(k)S(k)$, $W(k) = S_W^T(k)S_W(k)$, $V(k) = S^{\mathrm{T}}(k)\ S_v(k)$. In particular, $S(k) = \text{chol}(P(k))$, $S_W(k) = \text{chol}(W(k))$, $S_V(k) = \text{chol}(V(k))$, are the upper triangular Cholesky factorizations of $P(k)$, $W(k)$, and $V(k)$, respectively. The time updates for $k = 1, 2, \ldots$, are given by $x_+(k) = A(k-1)\hat{x}(k-1)$, $U(k) = \text{triu}(\text{qr}([S(k-1)A^{\mathrm{T}}(k-1); S_W(k-1)]))$, $P_{+s}(k) = U(k)(1{:}n;\ 1{:}n)$. The measurement updates are given by $P_+(k) = P_{+s}^{\mathrm{T}}(k)P_{+s}(k)$,/,$K(k) = P_+(k)C^+(k)[C(k)P_+(k)C^+(k) + V(k)]^{-1}$, $\hat{x}(k) + K(k)(y(k) - C(k)x_+(k))$, $Z(k) = \text{triu}(qr([S_v(k),\ 0_{mn}; P_{+s}(k)C^+(k), P_{+s}(k)]))$ and $S(k) = \text{Z}(k)(m+1{:}\ m+n, m+1{:}\ m+n)$.
3. **Square-root information filter (SRIF)**. In the SRIF form of the KF, we propagate the square root of the information matrix. Just as in the SRCF approach, as compared to the conventional covariance form of the KF, the SRIF approach, as compared to the SRIF approach, needs to use a lower dynamic range in the computations and obtain a more stable solution under finite precision computations. First, we denote $SI(k) = (\text{chol}(P(k)))^{-1}$, $SI_W(k) = (\text{chol}(W(k)))^{-1}$, and

$SI_V(k) = (\text{chol}(V(k)))^{-1}$. The time updates for $k = 1, 2, \ldots$, are given by $U(k) = \text{triu}(\text{qr}([SI_W(k-1), 0_{n\times n}, 0_{n\times 1}; SI(k-1)A^{-1}(k-1), SI(k-1)A^{-1}(k-1), b(k-1)]))$, $P_{+S}(K) = U(k)(n+1{:}2n, n+1{:}2n)$ and $b_+(k) = U(k)(n+1{:}2n, 2n+1)$. The measurement updates are given by $Z(k) = \text{triu}(\text{qr}([P_{+S}(k), b_+(k); SI_V(k)C(k), SI_V(k)y(k)]))SI(k) = Z(k)(1{:}n, 1{:}n)$, and $b(k) = Z(k)(1{:}n, n+1)$. At any iteration, $\hat{x}(k)$ and $P(k)$ are related to $b(k)$ and $SI(k)$ by $\hat{x}(k) = SI(k)b(k)$ and $P(k) = (SI^{\text{T}}(k)SI(k))^{-1}$.

9.5.5.3 Systolic Matrix Implementation of the KF Predictor

The covariance KF for the optimum state estimate $\hat{x}(k)$ includes the KF predictor $x_+(k)$ In particular, if we are only interested in $x_+(k)$ a relatively simple algorithm for $k = 1, 2, \ldots$, is given by $K(k) = P_+(k)C^{\text{T}}(k)[C(k)P_+(k)C^{\text{T}}(k) + V(k)]^{-1}$, $x_+(k+1) = A(k)\ x_+(k) + A(k)\ K(k)\ [y(k) - C(k)x_+(k)]$ and $P_+(k+1) = A(k)P_+(k)A^{\text{T}}(k) - A(k)K(k)C(k)P_+(k)A^{\text{T}}(k) + W(k)$. To start this KF prediction algorithm, we use $\hat{x}(0)$ and $P(0)$ to obtain $x_+(1) = A(0)\hat{x}(0)$ and $P_+(1) = A(0)P(0)A^{\text{T}}(0) + W(0)$. The above operations involve matrix inverse; matrix–matrix and matrix–vector multiplications; and matrix and vector additions. Fortunately, the matrix inversion of $\alpha = C(k)P_+(k)C^{\text{T}}(k) + V(k)$ can be approximated by the iteration of $\beta(i+1) = \beta(i)[2I - \alpha\ \beta(i)]$, $i = 1, \ldots, I$. Here, $\beta(i)$ is the ith iteration estimate of the inverse of the matrix α. While the preceding equation is not valid for arbitrary α and $\beta(i)$, for KF applications, we can use $I = 4$ because a good initial estimate $\beta(1)$ of the desired inverse is available from the previous step in the KF. Clearly, with the use of the above equation for the matrix inversion, all the operations needed in the KF predictor can be implemented on an orthogonal array using systolic matrix operations of the form $D = B \times A + C$, as shown in Figure 9.113.

The recursive algorithm of the KF predictor is decomposed as a sequence of matrix multiplications, as shown in Table 9.14. In step 1 the $n \times n$ matrix $P_+(k)$ and the $m \times n$ matrix $C^{\text{T}}(k)$ are denoted as B and A, respectively. The rows of B (starting from the n, $n-1, \ldots, 1$ row) are skewed and inputted to the $n \times n$ array starting at time 1. By time n (as shown in Figure 9.113), all the elements of the first column of B (i.e., $b_{n1}, \ldots, b_{11}$) are in the first column of the array. At time $n+1, \ldots, 2n-1$, elements of the second to nth columns of B are inputted to the array and remain there until the completion of the BA matrix multiplication. At time $n+1$, a_{11} enters $(1,1)$ cell and starts the BA process. At time $n+m$, a_{1m} enters the $(1,1)$ cell. Of course, additional times are needed for other elements in the second to the nth rows of A to enter the array. Further processing and propagation times are needed before all the elements of $D = BA = P_+(k)C^{\text{T}}(k)$ are outputted. However, in step 2, because B remains as $P_+(k)$, we do not need to input it again, but only append $A(k)$ (denote as $\tilde{A}$ in Figure 9.113) in the usual skewed manner after the previous $A = C^{\text{T}}(k)$. Thus, at time $n+m+1$, $\tilde{a}_{11}$ enters the (1, 1) cell. By time $n+m+n$, $\tilde{a}_{1n}$ enters the (1, 1) cell. Thus, step 1 takes $n+m$ time units, while step 2 takes only n time units. In step 3 m time units are needed to load $C(k)$ and m time units are needed to input $P_+(k)C^{\text{T}}(k)$, resulting in $2m$ time units. Steps 4 and 5 perform one iteration of the inverse approximation. In general, $I = 4$ iterations is adequate, and $16m$ time units are needed. Thus far, all the matrices and vectors are fed continuously into the array with no delay. However, in order to initiate step 13, the $(n, 1)$ component of $A(k) - A(k)K(k)C(k)$ is needed, but not available. Thus, at the end of step 11, an additional $(n-3)$ time units of delay must be provided to access this component. From Table 9.14, a total of $9n + 22m$ time units is needed to perform one complete KF prediction iteration.

9.5.5.4 Systolic KF Based on the Faddeev Algorithm

A form of KF based on mixed prediction error covariance $P_+(k)$ and information matrix $PI(k) = P^{-1}(k)$ updates can be obtained from the covariance KF algorithm. For $k = 1, 2, \ldots$, we have $x_+(k) = A(k-1)\hat{x}(k-1) + B(k-1)$, $P_+(k) = A(k-1)PI^{-1}(k-1)A^{\text{T}}(k-1) + W(k-1)$, $PI(k) = P_+^{-1}(k) + C^{\text{T}}(k)V^{-1}(k)C(k)K(k) = PI^{-1}(k)C^{\text{T}}(k)V^{-1}(k)$ and $\hat{x}(k) = x_+(k) + K(k)(y(k) - C(k)x_+(k))$. The algorithm starts with the given $\hat{x}(0)$ and $P(0)$, as usual. Because this algorithm requires the repeated use of matrix inversions for $(PI(k-1))$, $(P_+(k))^{-1}$, $(V(k))^{-1}$ as well as $P(k) = (PI(k))^{-1}$, the following "Faddeev algorithm" is suited for this approach.

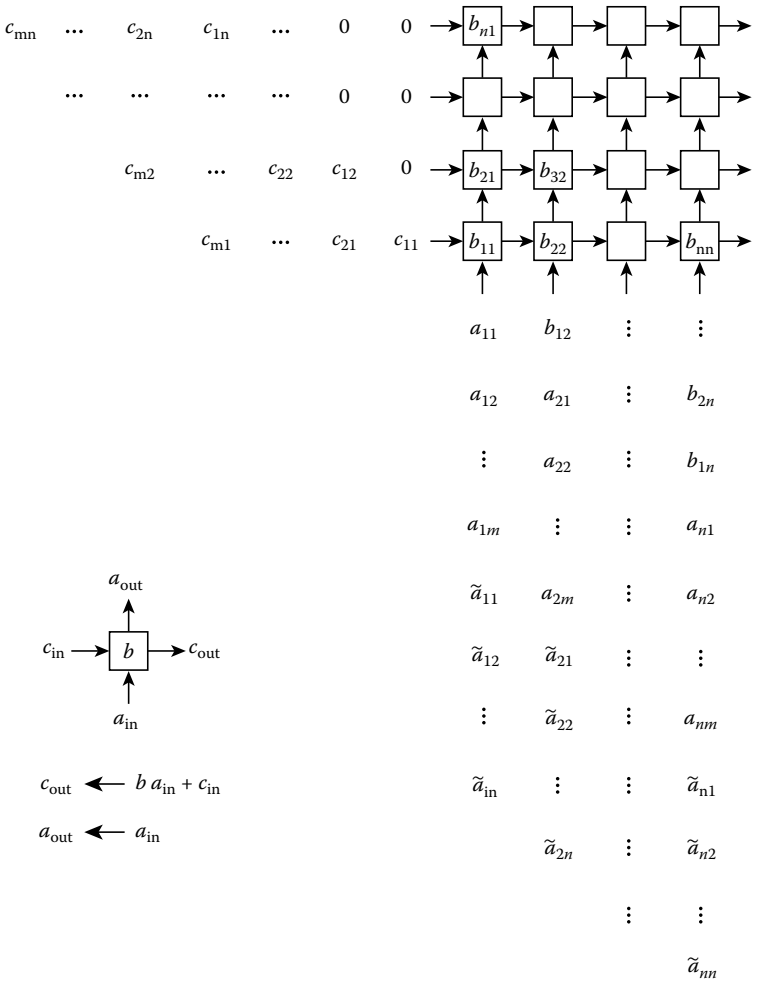


FIGURE 9.113 Systolic matrix multiplication and addition of $B \times A + C$.

TABLE 9.14 Systolic Matrix Operations of a KF Predictor

Step	B	A	C	D	Time
1	$P_+(k)$	$C^T(k)$	0	$P_+(k)C^T(k)$	$n+m$
2	$P_+(k)$	$A^T(k)$	0	$P_+(k)A^T(k)$	N
3	$C(k)$	$P_+(k)\ C^T(k)$	$V(k)$	$C(k)P_+(k)\ C^T(k)+V(k)=\alpha$	$2m$
4	α	$-\beta(i)$	$2I$	$2I-\alpha\beta(i)$	$2Im$
5	$\beta(i)$	$2I-\alpha\beta(i)$	0	$\beta(i+1)$	$2Im$
6	$P_+(k)\ C^{\mathrm{T}}(k)$	β	0	$K(k)$	$n+m$
7	$A(k)$	$K(k)$	0	$A(k)K(k)$	$n+m$
8	$A(k)$	$x_+(k)$	0	$A(k)x_+(k)$	1
9	$-C(k)$	$x_+(k)$	$y(k)$	$y(k)-C(k)x_+(k)$	$m+1$
10	$A(k)K)(k)$	$-C(k)$	$A(k)$	$A(k)-A(k)K(k)C(k)$	$2n$
11	$A(k)K(k)$	$y(k)-C(k)x_+(k)$	$A(k)x_+(k)$	$x_+(k+1)$	1
12	$n-3$				
13	$A(k)-A(k)K(k)C(k)$	$P_+(k)A^T(k)$	$W(k)$	$P_+(k)$	$2n$

Consider an $n \times n$ matrix A, an $n \times m$ matrix B, a $p \times n$ matrix C, and a $p \times m$ matrix D arranged in the form of a compound matrix $[A\ B; -C\ D]$. Consider a $p \times n$ matrix W multiplying $[A\ B]$ and added to $[-C\ D]$, resulting in $[A\ B; -C + WAD + WB]$. Assume W is chosen such that $-C + WA = 0$, or $W = CA^{-1}$. Then, we set $D + WB = D + CA^{-1}B$.

In particular, by picking $\{A, B, C, D\}$ appropriately, the basic matrix operations needed above can be obtained using the Faddeev algorithm. Some examples are given by

$$\begin{matrix} A & I \\ -I & 0 \end{matrix} \Rightarrow D + W\ B\ = A^{-1}$$

$$\begin{matrix} I & B \\ -C & 0 \end{matrix} \Rightarrow D + W\ B\ = C\ B$$

$$\begin{matrix} I & B \\ -C & D \end{matrix} \Rightarrow D + W\ B\ = D + C\ B$$

$$\begin{matrix} A & B \\ -I & 0 \end{matrix} \Rightarrow D + W\ B\ = A^{-1}B$$

A modified form of the previous Faddeev algorithm first triangularizes A with an orthogonal transformation Q, which is more desirable from the finite precision point of view. Then, the nullification of the lower left portion can be performed easily using the Gaussian elimination procedure. Specifically, applying a QRD, $Q[A\ B] = [R\ QB]$. Then, applying the appropriate W yields

$$\begin{bmatrix} R & QB \\ -C + W\ Q\ A & D + W\ Q\ B \end{bmatrix} = \begin{bmatrix} R & Q\ B \\ 0 & D + C\ A^{-1}B \end{bmatrix} \tag{9.24}$$

The preceding mixed prediction error covariance and information matrix KF algorithm can be reformulated as a sequence of Faddeev algorithm operations, as given in Table 9.15. The times needed to

TABLE 9.15 Faddeev Algorithm Solution to KF

Step	Compound Matrix		$D + WB$	Time
1	I $-A(k-1)$	$\hat{x}(k-1)$ $B(k-1)u(k-1)$	$x_+(k)$	$n+1$
2	$P^{-1}(k-1)$ $-A(k-1)$	$A^{\mathrm{T}}(k-1)$ $W(k-1)$	$P_+(k)$	$2n$
3	$V(k-1)$ $-C^{\mathrm{T}}(k)$	I 0	$C^{\mathrm{T}}(k)V^{-1}(k-1)$	$m+n$
4	$P_+(k)$ $-I$	I 0	$p_+^{-1}(k)$	$2n$
5	I $-C^{\mathrm{T}}(k)V^{-1}(k)$	$C(k)$ $P_+^{-1}(k)$	$P^{-1}(k)$	N
6	$P^{-1}(k)$ $-I$	$C^{\mathrm{T}(k+1)V^{-1}(k)}$ 0	$K(k)$	$2n$
7	I $C(k)$	$x_+(k)$ $y(k)$	$y(k) - C(k)_{x_+}(k)$	$m+1$
8	I $-K(k)$	$y(k) - C(k)x_+(k)$ $x_+(k)$	$\hat{x}(k)$	$m+1$

perform steps 2, 3, 4, and 6 are clearly just the sum of the lengths of the two matrices in the corresponding steps. Step 1 requires only n times unit to input the second row of matrices because $\hat{x}(k-1)$ is already located in the array from the previous iteration (step 8 output) and one time unit to output $x_+(k)$. Due to the form of $[-I\ 0]$ in step 4, $C(k)$ of step 5 can be inputted before the completion of $P_+^{-1}(k)$ in step 4. Thus, only n time units are needed in step 5. Similarly, $x_+(k)$ of step 7 can be inputted in step 6. Thus, we need only $m+1$ time units to input $[C(k)\ y(k)]$ and complete its operations. In step 8 only $m+1$ time units are needed as in step 1. Thus, a total of $9n+3m+3$ time units are needed for the Faddeev algorithm approach to the KF.

9.5.5.5 Other Forms of Systolic KF and Conclusions

While the operations of a KF can be expressed in many ways, only some of these algorithms are suited for systolic array implementations. For a KF problem with a state vector of dimension n and a measurement vector of dimension m, we have shown the systolic matrix–matrix multiplication implementation of the predictor form of the KF needs $9n+22m$ time steps for each iteration. A form of KF based on mixed update of prediction error covariance and information matrices is developed based on the Faddeev algorithm using matrix–matrix systolic array implementation. It has a total of $9n+3m+3$ time steps per iteration. A modified form of the SRIF algorithm can be implemented as a systolic array consisting of an upper rectangular array of $n(n+1)/2$ internal cells, and a lower n-dimensional triangular array of n boundary cells, and $(n-1)^2/2$ internal cells, plus a row of n internal cells, and $(n-1)$ delay cells. It has a total of n-boundary cells, $((n-1)^2+2n^2+2n)/2$ internal cells, and $(n-1)$ delay cells. Its throughput rate is $3n$ time steps per iteration. A modified form of the SRCF algorithm utilizing the Faddeev algorithm results in a modified SRCF form of a KF consisting of a trapezodial section, a linear section, and a triangular section systolic array. The total of these three sections needs $(n+m)$ boundary cells, n linear cells, and $((m-1)^2+2nm+(n-1)^2)/2$ internal cells. Its throughput rate is $3n+m+1$ time steps per iteration. The operations of both of these systolic KF are quite involved and detailed discussions are omitted here. In practice, in order to compare different systolic KFs, one needs to concern oneself not only with the hardware complexity and the throughput rate, but other factors involving the number of bits needed finite precision computations, data movement in the array, and I–O requirements as well.

9.5.6 Eigenvalue and SVDs

Results from linear algebra and matrix analysis have led to many powerful techniques for the solution of wide range of practical engineering and signal processing problems. Although known for many years, these mathematical tools have been considered too computationally demanding to be of any practical use, especially when the speed of calculation is an issue. Due to the lack of computational power, engineers had to content themselves with suboptimal methodologies of simpler implementation. Only recently, due to the advent of parallel/systolic computing algorithms, architectures, and technologies, have engineers employed these more sophisticated mathematical techniques. Among these techniques are the so-called eigenvalue decomposition (EVD) and the SVD. As an application of these methods, we consider the important problem of spatial filtering.

9.5.6.1 Motivation–Spatial Filtering Problem

Consider a linear array consisting of L sensors uniformly spaced with an adjacent distance d. A number M, $M<L$, of narrowband signals of center frequency f_0, impinging on the array. These signals arrive from M different spatial direction angles $\theta_1,\ldots,\theta_M$, relative to some reference direction. Each sensor is provided with a variable weight. The weighted sensor outputs are then collected and summed. The goal is to compute the set of weights to enhance the estimation of the desired signals arriving from directions $\theta_1,\ldots,\theta_M$.

In one class of beamformation problems, one sensor (sometimes referred to as main sensor) receives the desired signal perturbed by interference and noise. The remaining $L-1$ sensors (auxiliary sensors)

are mounted and aimed in such a way as to collect only the (uncorrelated) interference and noise components. In this scenario the main sensor gain is to be kept at a fixed value, while the auxiliary weights are adjusted in such a way as to cancel out as much perturbation as possible. Obviously, the only difference of this latter cast is that one of the weights (the one corresponding to the main sensor) is kept at a constant value of unity.

Let the output of the ith sensor, $i = 1, \ldots, L$, at discrete time $n = 0, 1, \ldots$, be given by

$$\bar{x}_i(n) = \Re\left\{ [x_i(n) + v_i(n)] e^{j(2\pi f_0 n)} \right\} \quad x_i(n) = a_i(n) \sum_{k=1}^{M} S_k(n) e^{j2\pi(i-1)d \sin\theta_k/\lambda}$$

where

$a_i(n)$ is the antenna gain at time n
S_k is the complex amplitude of the kth signal, inclusive of the initial phase
λ is the signal wavelength

The vectors $x_i(n)$ and $v_i(n)$ are analytic signal representations. The noise $v_i(n)$ is assumed to be uncorrelated white and Gaussian, of power σ_N^2. In order to avoid the ill effects of spatial aliasing, let us also assume that $d \leq \lambda/2$. The outputs of the sensor array for times $n = 0, 1, \ldots, N$, can be collected in matrix form as follows:

$$\underbrace{X}_{N\times L} = \underbrace{S}_{N\times M} = \underbrace{A}_{M\times L} + \underbrace{V}_{N\times L}$$

The matrix A is referred to as the "steering matrix." In the case in which $a_i(n) = 1$ for all i, the matrix A is Vandermonde and full rank, and its kth row can be expressed as $A(\theta) = (1, e^{j2\pi d \sin\theta/\lambda}, \ldots, e^{j2\pi(L-I)d \sin\theta_k/\lambda})$.

The data correlation matrix, $R_X = E\{X^H X\}$, where $E(\cdot)$ is the ensemble average operator, is equal to $R_X = A^H R_s A + \sigma_N^2 I, R_s = \mathrm{E}\{S^H S\}$. We note:

1. The matrix R_s has rank M by definition as does the matrix $A^H R_s A$.
2. The rows of A are in the range space of R_X.
3. The value σ_N^2 is an eigenvalue of R_X with multiplicity $L - M$, given the det $(R_X - \sigma_N^2 I) = 0$, and the rank of $A^H R_s A$ is M.

The EVD of R_X can therefore be written as $R_X = V_S \Lambda_S V_S^H + \sigma_N^2 V_N V_N^H$, where V_S is $L \times M$, V_N is $L \times (L - M)$, $V_S^H V_S = I$ and $V_N^H V_N = 1$, and $V_S^H V_N = 0$. Moreover, we have that $A V_N = 0$. Let $A(\theta)$ be a generic steering vector, defined as $A(\theta) + (1, e^{j2\pi d \sin\theta/\lambda}, \ldots, e^{j2\pi(L-I)d \sin\theta/\lambda})$. Then, the function $\pm(\theta) = 1/|A(\theta) V_N|^2$ has M poles at the angles $\theta = \theta_k k = 1, \ldots, M$. Alternatively, any linear combination, w, of the columns of V_N is such that $E\{\|Xw\|_2\} = \min_z E\{\|Xz\|_2\} = \sigma_N$. In other words, the signals impinging from angular directions $\theta_1, \ldots, \theta_M$ are totally canceled out in the system output.

The desired weighting vector for our spatial filtering problem can consequently be expressed as $w = V_N p$, $p = [p_1, \ldots, p_{L-M}]^{\mathrm{T}}$, for any nonzero vector p. From the above discussion, we see that the solution to the original spatial filtering problem can be obtained from the EVD of the correlation matrix $R_X = E\{X^H X\}$. In practice the sample correlation matrix $\hat{R}_X$ is used instead, where the ensemble average is replaced by a suitable temporal average.

The computation of the covariance matrix implies the computation of the matrix product $X^H X$. Some small elements in X are then squared and the magnitude of the resulting element can become comparable or smaller than the machine precision. Rounding errors can often impair and severely degrade the computed solution. In these cases it is better to calculate the desired quantities (correlation eigenvalues and eigenvectors) directly from the data matrix X using the SVD technique as considered next.

9.5.6.2 EVD of a Symmetric Matrix

Consider an $L \times L$ real symmetric matrix $A = A^{\mathrm{T}}$. In the previous spatial filtering example, $A = R_X$. Let

$$G(i,j,\theta) = \begin{array}{c} \\ i \\ \\ \\ \\ j \\ \\ \end{array} \begin{array}{c} \begin{array}{ccccc} i & & j & & \end{array} \\ \begin{pmatrix} I_{i-1} & & & & \\ & c & & s & \\ & & I_{j-i-1} & & \\ & -s & & c & \\ & & & & I_{L-j} \end{pmatrix} \end{array}$$

be an "orthogonal Givens rotation matrix," where $c = \cos\theta$ and $s = \sin\theta$. Pre- or postmultiplication of A by G leaves A unchanged, except for rows (columns) i and j, which are replaced by a linear combination of old rows (columns) i and j. A "Jacobi rotation" is obtained by simultaneous pre- and postmultiplication of a matrix by a Givens rotation matrix, as given by $G(i, j, \theta)^{\mathrm{T}} A G(i, j, \theta)$, where θ is usually chosen in order to zero out the (i, j) and (j, i) entries of A.

The matrix A can be driven toward diagonal form by iteratively applying Jacobi rotations, as given by $A_0 \leftarrow A$, $A_{k+1} \leftarrow G_k^T A G_k$, where G_k is a Givens rotation matrix. A "sweep" is obtained by applying $L(L-1)/2$ Jacobi rotations, each nullifying a different pair of off-diagonal elements, according to a prespecified order. Given the matrix $A_k = (a_{pq}^{(k)})$ at the kth iteration, and a pair of indices (i, j), the value of $\tan\theta$ can be obtained from the following equations:

$$u = \frac{a_{jj}^{(k)} - a_{ii}^{(k)}}{2a_{ij}^{(k)}}, \quad \tan\theta = \frac{sign(u)}{|u| + \sqrt{1+u^2}} \tag{9.25}$$

It is possible to demonstrate that each Jacobi rotation reduces the matrix off-norm. The matrix A_k indeed tends to diagonal form and for all practical purposes it reaches it after $\mathbb{O}(\log L)$ sweeps.

The matrix V of eigenvectors is obtained by applying the same rotations to a matrix initialized to the identity, as follows: $V_0 \leftarrow I$, $V_{k+1} \leftarrow V_k G_k$. A two-dimensional systolic array implementation of the previous algorithm is shown in Figure 9.114, for the case $L = 8$. At the beginning of iteration k, processor P_{ij} contains elements

$$\begin{pmatrix} a_{2i-1,2j-1}^{(k)} & a_{2i-1,2j}^{(k)} \\ a_{2i,2j-1}^{(k)} & a_{2i,2j}^{(k)} \end{pmatrix} \quad i,j = 1,2,\ldots,L/2,$$

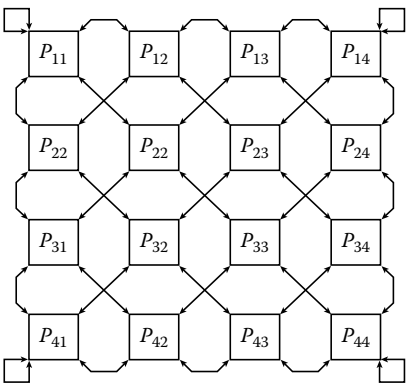


FIGURE 9.114 Systolic array for an EVD of a symmetric matrix based on Jacobi rotations.

The diagonal processors compute the rotation parameters and apply the rotation to the four entries they store. Subsequently, they propagate the rotation parameters horizontally and vertically to their neighbors, which, upon receiving them, apply the corresponding rotation to their stored entries. After the rotation is applied, each processor swaps its entries with its four neighbors along the diagonal connections. The correct data movement at the edges of the array is also shown in Figure 9.114. A correct scheduling of operations requires that each processor be idle for two out of three time steps, which translates into an "efficiency" of 33%. Each sweep takes 3 $(L-1)$ time steps, and the number of sweeps can be chosen on the order of $\log L$.

9.5.6.3 SVD of a Rectangular Matrix via the Hestenes Algorithm

Consider an $N \times L$ real matrix A, $N \geq L$. Its SVD can be written as follows:

$$\underbrace{A}_{N\times L} = \underbrace{U}_{N\times L}\,\underbrace{\Sigma}_{L\times L}\,\underbrace{V^{\mathrm{T}}}_{L\times L}$$

where U and V have orthonormal columns. The matrix $\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_L)$ is the diagonal matrix of singular values, where $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_L \geq 0$. Consider the following recursion $A_0 \leftarrow A$, $A_{k+1} \leftarrow A_k\, G_k$, where the Givens rotations are chosen not to zero out entries of A_k, but to orthogonalize pairs of its columns. A sweep is now defined as a sequence of Givens rotations that orthogonalize all $\binom{L}{2}$ pairs of columns exactly once. Observe the similarity with the algorithm described previously for the calculation of eigenvalues and eigenvectors. If $G(i, j, \theta)$ is the Givens rotation which orthogonalizes columns i and j of A, then $G(i, j, \theta)^{\mathrm{T}} M G(i, j, \theta)$ is the Jacobi rotation that zeroes out the entries (i, j) and (j, i) of $M = A^{\mathrm{T}}A$. A sweep (as defined here) of rotations applied to the rectangular matrix A corresponds exactly to a sweep (as defined earlier) of rotations applied to the symmetric matrix M.

At any time step, the original matrix A can be expressed as follows:

$$A = A_k V_k^T, \quad V_k = \prod_{i=1}^{k} G_k$$

where V_k has orthonormal columns for any k (by definition of Givens rotations). After a number of sweeps (on the order of $\log L$) the matrix A_k approaches a matrix, W, of orthogonal columns, $A_k \rightarrow W$, $V_k \rightarrow V$. If σ_i is the norm of the ith column of W, $i = 1, \ldots, L$, then we have $W = U\ \mathrm{diag}\ (\sigma_1, \ldots, \sigma_L)$, $A = U\Sigma V^{\mathrm{T}}$.

This SVD approach based on the Hestenes algorithm can be realized on a Brent–Luk [58] linear systolic array, as shown in Figure 9.115, for the case $L = 8$. Each processor stores a pair of columns; in particular, the procedure starts by storing columns 2 $k-1$ and $2k$ in processor P_k. Each processor computes the rotation parameters which orthogonalize the pair of columns.

Let x and z be the two stored columns. Let ξ and ζ be their norms, and $\acute{\eta}$ be their inner product. Then the value of $\tan\theta$ from

$$u = \frac{\zeta - \xi}{2\eta}, \quad \tan\theta = \frac{\mathrm{sign}(u)}{|u| + \sqrt{1 + u^2}}$$

After applying the computed rotation, each processor swaps its pair of columns with the two neighboring processors along the connections shown in Figure 9.115. The column indices stored in each processor at the different steps of a single sweep are given in Table 9.16. Note that all the $\binom{L}{2}$ pairs of indices are generated by using the Brent–Luk scheme. The stopping criteria can be set in advance. A possible criterion is by inspecting the magnitude of the rotating angles. When they are all in absolute value below

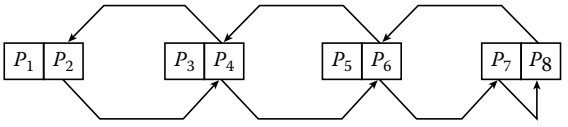


FIGURE 9.115 Linear systolic array for SVD of a rectangular matrix based on the Hestenes algorithm.

TABLE 9.16 Movement of Matrix Columns during One Sweep ($L = 8$)

P_1	P_2	P_3	P_4
(1, 2)	(3, 4)	(5, 6)	(7, 8)
(1, 4)	(2, 6)	(3, 8)	(5, 7)
(1, 6)	(4, 8)	(2, 7)	(3, 5)
(1, 8)	(6, 7)	(4, 5)	(2, 3)
(1, 7)	(8, 5)	(6, 3)	(4, 2)
(1, 5)	(7, 3)	(8, 2)	(6, 4)
(1, 3)	(5, 2)	(7, 4)	(8, 6)
(1, 2)	(3, 4)	(5, 6)	(7, 8)

a given threshold, then the algorithm can stop. More commonly, a predetermined number of sweeps is chosen ahead of time. Observation shows that a number of sweeps on the order of log L is sufficient for convergence.

9.5.6.4 SVD of a Rectangular Nonsymmetric Matrix

9.5.6.4.1 Via the Jacobi Algorithm

The SVD algorithm described in the previous section has the drawback of a somewhat complicated updating procedure. In many signal processing applications, continuous updating of the matrix decomposition as new samples are appended to the data matrix is required. Such problems occur in spectral analysis, direction-of-arrival estimation, beam forming, etc. An efficient updating procedure for the SVD of rectangular matrices of growing row size is given by the algorithm described in this section, based on the succession of two basic operations: a QR updating step, followed by a rediagonalization operation. This algorithm is otherwise known as a version of the Kogbetliantz algorithm for triangular matrices.

Given the $m \times L$ data matrix at time m, $A_m = [a_1, \ldots, a_m]^T$, where a_i, $i = 1, \ldots, m$ are the rows of A_m, one defines the exponentially weighted matrix $B_m(\beta)\, A_m$, where $B_m(\beta)$ is the diagonal forgetting matrix $B_m(\beta) + \text{diag}(\beta^{m-1}, \beta^{m-2}, \ldots, \beta, 1)$, and $0 < \beta \leq 1$ is the forgetting factor. The updating problem is to determine the SVD of the updated weighted matrix $B_{m+1}(\beta)\, A_{m+1}$, given the SVD at time m,

$$B_m(\beta) A_m = U_m \sum_m V_m$$

Often only the singular values and right singular vectors are of interest. This is fortunate because the left singular matrix grows in size as time increases, while the sizes of Σ_m and V_m remain unchanged.

The algorithm can be summarized as follows. Given the matrices V_m and Σ_m and the new data sample x_{m+1},

$$\sum\nolimits'_m \leftarrow \begin{pmatrix} \beta & \sum \\ & m \\ x_{m+1} & V_m \end{pmatrix} V'_m \leftarrow V_m,$$

the QR updating step

$$\begin{pmatrix} \sum'_m \\ 0 \end{pmatrix} \leftarrow Q_{m+1} \sum\nolimits'_m$$

the rediagonalization using permuted Jacobi rotations

for $k = 1, \ldots, 1$,
  for $i = 1, \ldots, n-1$, $j = i+1$,
    $\Sigma'_m \leftarrow \prod_{ij} G(i, j, \theta)\Sigma'_m G(i, j, \varphi)^{\mathrm{T}} \prod_{ij}$,
    $V'_m \leftarrow V'_m G(i, j, \varphi)^{T} \prod_{ij}$.
end
end
    $\Sigma_{m+1} \leftarrow \Sigma'_m$,
    $V_{m+1} \leftarrow V'_m$.

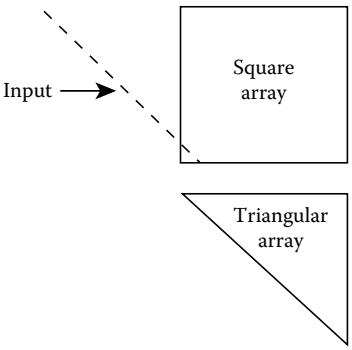


FIGURE 9.116 Two-dimensional systolic array for updating the Jacobi SVD algorithm for a nonsymmetric matrix.

In the preceding algorithm, the parameter l determines both the number of computations between subsequent updates and the estimation accuracy at the end of each update step. When l is chosen equal to the problem order L, then one complete sweep is performed. In practice, the l parameter can be chosen as high as $\sim 10L$ (usually for block computations or slow updates) and as small as 1 (for very high updating rates, with understandable degradation of estimation accuracy). The matrices Σ_m and Σ'_m are upper triangular at all times. This is ensured by the application of the permuted left and right Givens rotations in the rediagonalization step. After the application of any Jacobi rotation, the rotated rows and columns are subsequently permuted. This expedient not only preserves the upper triangularity of the Σ matrices, but also makes it possible for the rotations to be generated on the diagonal and propagated along physically adjacent pairs of rows and columns. All these features make this algorithm a very attractive candidate for a systolic implementation.

A schematic diagram of the systolic array proposed by Moonen et al. [63] is shown in Figure 9.116, where the triangular array stores the matrices Σ_m and Σ'_m, for all m, and the square array stores the V-matrix. The incoming data samples, x_{m+1}, are input into the V-array, where the vector-by-matrix multiplication $x_{m+1}\ V_m$ is performed. The output is subsequently fed into the triangular array. As it propagates through the array, the QR updating step is carried on: left rotations are generated in the diagonal elements of the array and propagated through the corresponding rows of the Σ matrix. One does not need to wait for the completion of the QR step to start performing the Jacobi rotations associated with the diagonalization step. It is known that the two operations (QR update and diagonalization) can be interleaved without compromising the final result. The parameters relative to the left rotations are, as before, propagated along the rows of the triangular matrix, while the right rotation parameters move along the columns of Σ_m, and are passed on to the V-array. Due to the continual modification of the matrix V_m caused by these right rotations, and because of the use of finite precision arithmetic, the computed right singular matrix may deviate from orthogonality. It is also known that in a realistic environment the norm of $V_m V_m^{\mathrm{T}} - I$ grows linearly with m. Reorthogonalization procedures must therefore be included in the overall scheme. A complicated reorthogonalization procedure based on left rotations, which interleaves with the other operations, was described in Moonen et al. [63]. An improved reorthogonalization algorithm was proposed by Vanpoucke and Moonen [66], where the matrix V_m is stored in parametrized form, thereby guaranteeing orthogonality at all times. The resulting triangular array and its modes of operation were also described.

References

1. A. Darte and J. M. Delosme, Partitioning for array processors, Tech. Rep. LIP-IMAG 90–23, Laboratoire de l'Informatique du Parallélisme, Ecole Supérieure de Lyon, Oct. 1990.
2. H. T. Kung and C. E. Leiserson, Systolic arrays (for VLSI), in *Introduction to VLSI Systems*, C. A. Meads and L. A. Conway, Eds. Reading, MA: Addison-Wesley, 1980, Chapter 8.3.

3. H. T. Kung, Why systolic architecture? *Computer*, 15, 37–46, 1982.
4. S. Y. Kung, K. S. Arun, R. J. Gal-Elzer, and D. V. B. Rao, Wavefront array processor: Language, architecture, and applications, *IEEE Trans. Comput*., 31, 1054–1066, 1982.
5. S. Y. Kung, *VLSI Array Processing*, Englewood Cliffs, NJ: Prentice Hall, 1988.
6. F. Lorenzelli, Systolic mapping with partitioning and computationally intensive algorithms for signal processing, PhD thesis, University of California, Los Angeles, 1993.
7. D. I. Moldovan, On the analysis and synthesis of VSLI algorithms, *IEEE Trans. Comput*., 31, 1121–1126, 1982.
8. M. Newman, *Integral Matrices*, New York: Academic Press, 1972.
9. P. Quinton, The systematic design of systolic arrays, IRISA Tech. Report 193, April 1983.
10. S. Rao, Regular iterative algorithms and their implementations on processor arrays, PhD thesis, Stanford, CA: Stanford University, 1985.
11. E. Angelidis and J. E. Diamessis, A novel method for designing FIR filters with nonuniform frequency samples, *IEEE Trans. Signal Process*., 42, 259–267, 1994.
12. S. Chanekar, S. Tantaratana, and L. E. Franks, Multiplier-free IIR filter realization using periodically time-variable state-space structures, I and II, *IEEE Trans*, *Signal Process*., 42, 1008–1027, 1994.
13. L. A. Ferrari and P. V Sankar, Minimum complexity FIR filters and sparse systolic arrays, *IEEE Trans. Comput*., 37, 760–764, 1988.
14. A. Jayadeva, A new systolic design for digital IIR filters, *IEEE Trans. Circuits Syst*., 37, 653–654, 1990.
15. S. C. Knowles, J. G. McWhirter, R. F. Woods, and J. V. McCanny, Bit-level systolic architectures for high performance IIR filtering, *J. VLSI Signal Process*., 1, 9–24, 1989.
16. S. M. Lei and K. Yao, A class of systolizable IIR digital filters and its design for proper scaling and minimum output roundoff noise, *IEEE Trans. Circuits Syst*., 37, 1217–1230, 1990.
17. S. M. Lei and K. Yao, Efficient systolic array implementation of IIR digital filtering, *IEEE Trans. Circuits Syst*., 39, 581–584, 1992.
18. H. H Loomis Jr. and B. Sinha, High-speed recursive digital filter, *Circuits*, *Syst. Signal Process*., 3, 267–294, 1984.
19. K. K. Parhi and D. G. Messerschmitt, Pipeline interleaving and parallelism in recursive digital filters. I. Pipelining using scattered look-ahead and decomposition, *IEEE Trans. Acoust. Speech Signal Process*, 37, 1099–1117, 1989.
20. N. R. Shanbhag and K. K. Parhi, *Pipelined Adaptive Digital Filters*, Boston, MA: Kluwer Academic, 1994.
21. R. F. Woods and J. V. McCanny, Design of high performance IIR digital filter chip, *IEEE Proc. E*, *Comput. Digital Tech*., 139, 195–202, 1992.
22. C. W. Wu and J.-C. Wang. Testable design of bit-level systolic block FIR filters, *Proc. IEEE Int. Symp. Circuits Syst*., San Diego, CA, pp. 1129–1132, 1992.
23. R. Wyrzykowski and S. Ovramenko, Flexible systolic architecture for VLSI FIR filters, *IEEE Proc. E*, *Comput. Digital Tech*., 139, 170–172, 1992.
24. L. W. Chang and J. H. Lin, A bit-level systolic array for median filter, *IEEE Trans. Signal Process*., 40, 2079–2083, 1992.
25. J. M. Delosme, Bit-level systolic array for real symmetric and Hermitian eigenvalue problems, *J. VLSI Signal Process*., 4, 69–88, 1992.
26. R. A. Evans, J. V. McCanny, J. G. McWhirter, A. Wood, and K. W. Wood, A CMOS implementation of a systolic multibit convolver chip, *Proc. VLSI*, 227–235, 1983.
27. G. Fettweis and H. Meyr, High-rate Viterbi processor: A systolic array solution, *IEEE J. Sel. Topics Commun*., 8, 1520–1534, 1990.
28. S. C. Knowles, J. G. McWhirter, R. F. Woods, and J. V. McCanny, Bit-level systolic architectures for high performance IIR filtering, *J. VLSI Signal Process*., 1, 9–24, 1989.
29. J. V. McCanny and J. G. McWhirter, On the implementation of signal processing function using one bit systolic arrays, *Elect. Lett*., 18, 241–243, 1982.

30. J. V. McCanny, J. G. McWhirter, and S. Y. Kung, The use of data dependence graphs in the design of bit-level systolic arrays, *IEEE Trans. Acoust., Speech Process.*, 38, 787–793, 1990.
31. J. V. McCanny, R. F. Woods, and M. Yan, Systolic arrays for high-performance digital signal processing, in *Digital Signal Processing: Principles, Device, and Applications*, N. B. Jones and J. D. M. Watson, Eds, New York: Peter Peregrinus, 1990, pp. 276–302.
32. C. L. Wang, An efficient and flexible bit-level systolic array for inner product computation, *J. Chin. Inst. Eng.*, 41, 567–576, 1991.
33. C. W. Wu, Bit-level pipelined 2-D digital filters image processing, *IEEE Trans. Circuits Syst. Video Technol.*, 1, 22–34, 1991.
34. M. G. Bellanger and P. A. Regalia, The FLS-QR algorithm for adaptive filtering: The case of multi-channel signal, *Signal Process*, 22, 115–126, 1991.
35. J. M. Cioffi, The fast adaptive ROTOR'S RLS algorithm, *IEEE Trans. Acoust. Speech Signal Process.*, 38, 631–653, 1990.
36. W. M. Gentleman and H. T. Kung, Matrix triangularization by systolic arrays, *Proc. SPIE, Real-Time Signal Process*, 298, 298–303, 1981.
37. S. Haykin, *Adaptive Filter Theory*, 2nd ed., Englewood Cliffs, NJ: Prentice Hall, 1991.
38. F. Ling and J. G. Proakis, A recursive modified Gram-Schmidt algorithm with applications to least-squares estimation and adaptive filtering, *IEEE Trans. Acoust. Speech Signal Process*, 34, 829–836, 1986.
39. K. J. R. Liu, S. F. Hsieh, K. Yao, and C. T. Chiu, Dynamic range, stability, and fault tolerant capability of finite-precision RLS systolic array based on Givens rotations, *IEEE Trans. Circuits Sys.*, June, 625–636, 1991.
40. K. J. R. Liu, S. F. Hsieh, and K. Yao, Systolic block householder transformation for RLS alogrithm with two-level pipelined implementation, *IEEE Trans. Signal Process.*, 40, 946–958, 1992.
41. J. G. McWhirter, Recursive least-squares minimization using a systolic array, *Proc. SPIE, Real-Time Signal Process. VI*, 431, 105–112, 1983.
42. J. G. McWhirter, Algorithm engineering in adaptive signal processing, *IEEE Proc. F.*, 139, 226–232, 1992.
43. B. Yang and J. F. Böhme, Rotation-based RLS algorithms: Unified derivations, numerical properties, and parallel implementations, *IEEE Trans. Acoust. Speech Signal Process.*, 40, 1151–1167, 1992.
44. M. J. Chen and K. Yao, On realization of least-squares estimation and Kalman filtering by systolic arrays, in *Systolic Arrays*, W. Moore, A. McCabe, and R. Urquhart, Eds. Bristol, U.K.: Adam Hilger, 1986, pp. 161–170.
45. F. Gaston, G. Irwin, and J. McWhirter, Systolic square root covariance Kalman filtering, *J. VLSI Signal Process.*, 2, 37–49. 1990.
46. J. H. Graham and T. F. Kadela, Parallel algorithm architectures for optimal state estimation, *IEEE Trans. Comput.*, 34, 1061–1068, 1985.
47. R. E. Kalman, A new approach to linear filtering and prediction problems, *J. Basic Eng.*, 82, 35–45, 1960.
48. P. G. Kaminiski, Discrete square root filtering: A survey of current techniques, *IEEE Trans. Autom. Control*, 16, 727–735, 1971.
49. S. Y. Kung and J. N. Hwang, Systolic array design for Kalman filtering, *IEEE Trans. Signal Process.*, 39, 171–182, 1991.
50. R. A. Lincoln and K. Yao, Efficient systolic Kalman filtering design by dependence graph mapping, in *VLSI Signal Processing III*, R. W. Brodersen and H. S. Moscovitz, Eds., New York: IEEE Press, 1988, pp. 396–407.
51. J. G. Nash and S. Hansen, Modified Faddeev algorithm for matrix manipulation, *Proc. SPIE*, 495, 39–46, 1984.
52. C. C. Paige and M. A. Saunders, Least squares estimation of discrete linear dynamic systems using orthogonal transformation, *SIAM J. Numer. Anal.*, 14, 180–193, 1977.

53. G. M. Papadourakis and F. J. Taylor, Implementation of Kalman filters using systolic arrays, *Proc. Int. Conf. Acoust*., *Speech and Signal Process*., Denver, CO, pp. 783–786, 1987.
54. P. Rao and M. A. Bayoumi, An algorithm specific VLSI parallel architecture for Kalman filter, in *VLSI Signal Processing IV*, H. S. Moscovitz, K. Yao, and R. Jain, Eds., New York: IEEE Press, 1991, pp. 264–273.
55. T. Y. Sung and Y. H. Hu, Parallel implementation of the Kalman filter, *IEEE Trans. Aero. Electr. Syst*., 23, 215–224, 1987.
56. H. Yeh, Systolic implementation of Kalman filters, *IEEE Trans. Acoust*., *Speech*, *Signal Process*., 36, 1514–1517, 1988.
57. E. Biglieri and K. Yao, Some properties of singular value decomposition and their application to signal processing, S*ignal Process*., 18, 277–289, 1989.
58. R. Brent and F. T. Luk, The solution of singular-value and symmetric eigenvalue problems on multiprocessor arrays, *SIAM J. Sci. Stat. Comput*., 6, 69–84, 1985.
59. G. H. Golub and C. F. Van Loan, *Matrix Computations*, 2nd ed., Baltimore, MD: Johns Hopkins University Press, 1989.
60. S. Haykin, *Adaptive Filter Theory*, 2nd ed., Englewood Cliffs, NJ: Prentice Hall, 1991.
61. M. R. Hestenes, Inversion of matrices by biorthogonalization and related results, *J. Soc. Ind. Appl. Math*., 6, 51–90, 1958.
62. F. T. Luk, A triangular processor array for computing singular values, *Linear Algebra Appl*., 77, 259–273, 1986.
63. M. Moonen, P. Van Dooren, and J. Vandewalle, A systolic array for SVD updating, *SIAM J. Matrix Anal. Appl*., 14, 353–371, 1993.
64. R. O. Schmidt, A signal subspace approach to multiple emitter location and spectral estimation,'' PhD thesis, Stanford, CA: Stanford University, 1981.
65. G. W. Stewart, A Jacobi-like algorithm for computing the Schur decomposition of a nonhermitian matrix, *SIAM J. Sci. Stat. Comput*., 6, 853–864, 1985.
66. F. Vanpoucke and M. Moonen, Numerically stable Jacobi array for parallel subspace tracking, *Proc. SPIE*, 2296, 403–412, 1994.
67. F. D. Van Veen and K. M. Buckley, Beamforming: A versatile approach to spatial filtering, *IEEE ASSP Mag*., 5, 4–24, 1988.

10

Data Converters

10.1 Digital-to-Analog Converters **10**-1
Introduction • Basic Converter Design Issues • Converter Architectures • Techniques for High-Resolution DACs • Sources of Conversion Errors • Low-Spurious DAC Design Examples
References **10**-28
10.2 Analog-to-Digital Converters **10**-29
Introduction • Nyquist Rate Converters • Oversampled Converters
Acknowledgment **10**-57
References **10**-57
Further Information **10**-58

Bang-Sup Song
University of California, San Diego

Ramesh Harjani
University of Minnesota

10.1 Digital-to-Analog Converters

Bang-Sup Song

10.1.1 Introduction

Digital-to-analog converters (DACs), referred to as decoders in communications terms, are devices by which digital processors communicate with the analog world. Although DACs are used as key elements in analog-to-digital converters (ADCs), they find numerous applications as stand-alone devices from CRT display systems and voice/music synthesizers to automatic test systems, waveform generators, digitally controlled attenuators, process control actuators, and digital transmitters in modern digital communications systems.

The basic function of the DAC is the conversion of input digital numbers into analog waveforms. An N-bit DAC provides a discrete analog output level, either voltage or current, for every level of 2^N digital words, $\{D_i;\ i = 0, 1, 2, \ldots, 2^N - 1\}$, that is applied to the input. Therefore, an ideal voltage DAC generates 2^N discrete analog output voltages for digital inputs varying from $000\ldots00$ to $111\ldots11$ as illustrated in Figure 10.1 for the 4 bit example. The output has a one-to-one correspondence with the input

$$V_{\text{out}}(D_i) = V_{\text{ref}}\left(\frac{b_N}{2} + \frac{b_{N-1}}{2^2} + \cdots + \frac{b_2}{2^{N-1}} + \frac{b_1}{2^N}\right) \tag{10.1}$$

where V_{ref} is a reference voltage setting the output range of the DAC and $b_N, b_{N-1}, \ldots, b_1$ is the binary representation of the input digital word D_i. In the unipolar case, as shown, the reference point is 0 when the digital input D_0 is $000\ldots00$, but in bipolar or differential DACs, the reference point is the midpoint of

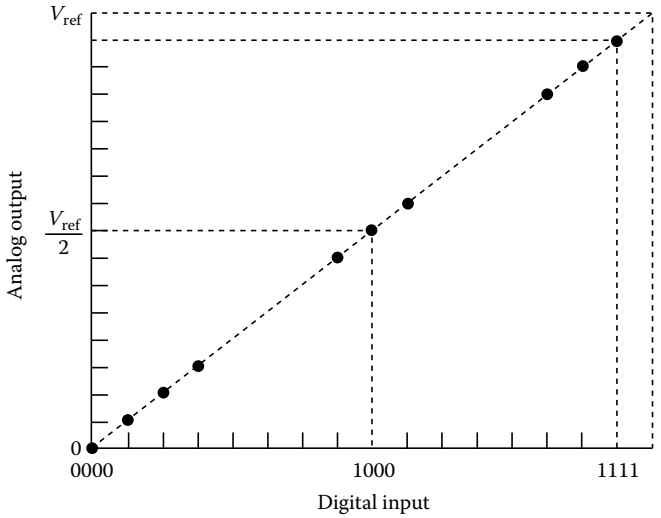


FIGURE 10.1 Transfer characteristics of a unipolar DAC.

the full scale when the digital input is 100 . . . 00, and the range is defined from $-V_{ref}/2$ to $V_{ref}/2$. Although purely current-output DACs are possible, voltage-output DACs are common in most applications.

10.1.1.1 Signal-to-Noise Ratio and Dynamic Range

The resolution is a term used to describe a minimum voltage or current that a DAC can resolve. The fundamental limit of a DAC is the quantization noise due to the finite resolution of the DAC. If the input digital word is N bits long, the minimum step that a DAC can resolve is $V_{ref}/2^N$. If output voltages are reproduced with this minimum step of uncertainty, an ideal DAC should have a maximum signal-to-noise ratio (SNR) of

$$\text{SNR} = \frac{3}{2}2^{2N} \approx 6N + 1.8 \text{ (dB)} \tag{10.2}$$

where SNR is defined as the power ratio of the maximum signal to the inband uncorrelated noise. For example, an ideal 16 bit DAC has an SNR of about 97.8 dB. The spectrum of the quantization noise is evenly distributed up to the Nyquist bandwidth (half the sampling frequency). Therefore, this inband quantization noise decreases by 3 dB when the oversampling ratio (OSR) is doubled. This implies that when oversampled, the SNR within the signal band can be made lower than the quantization noise limited by Equation 10.2.

The resolution of a DAC is usually characterized in terms of SNR, but the SNR accounts only for the uncorrelated noise. The real noise performance is better represented by TSNR, which is the ratio of the signal power to the total inband noise, including harmonic distortion (HD). Also, a slightly different term is often used in place of the SNR. The useful signal range or dynamic range is defined as the power ratio of the maximum signal to the minimum signal. The minimum signal is defined as the smallest input, for which the TSNR is 0 dB, while the maximum signal is the full-scale input. Therefore, the SNR of nonideal DACs can be lower than the ideal dynamic range because the noise floor can be higher with a large signal present. In practice, DACs are limited not only by the quantization noise, but also by nonideal factors such as noises from circuit components, power supply coupling, noisy substrate, timing jitter, insufficient settling, and nonlinearity.

10.1.2 Basic Converter Design Issues

The performance of a DAC can be specified in terms of its linearity, monotonicity, and conversion speed. In most conventional DACs, except for the oversampling DACs the linearity and monotonicity are limited by how accurately the reference voltage/current is divided using passive/active components.

10.1.2.1 Linearity

Differential nonlinearity (DNL). The output range of an N-bit DAC is equally divided into 2^N small units, as shown in Figure 10.1, and one least significant bit (LSB) change in the input digital word makes the analog output voltage change by $V_{\text{ref}}/2^N$. The DNL is a measure of the deviation of the actual DAC output voltage step from this ideal voltage step for 1 LSB. The DNL is defined as

$$\text{DNL} = \frac{V_{\text{out}}(D_{i+1}) - V_{\text{out}}(D_i) - V_{\text{ref}}/2^N}{V_{\text{ref}}/2^N} \quad \text{for } i = 0, 1, \ldots, 2^N - 2 \text{ (LSB)} \tag{10.3}$$

and the largest positive and negative numbers are usually quoted to specify the static performance of a DAC.

Integral nonlinearity (INL). The overall linearity of a DAC can be specified in terms of the INL, which is a measure of deviation of the actual DAC output voltage from the ideal straight line drawn between two endpoints, 0 and V_{ref}. Because the ideal output is $i \times V_{\text{ref}}/2^N$ for any digital input D_i, the INL is defined as

$$\text{INL} = \frac{V_{\text{out}}(D_i) - i \times V_{\text{ref}}/2^N}{V_{\text{ref}}/2^N} \quad \text{for } i = 0, 1, \ldots, 2^N - 1 \text{ (LSB)} \tag{10.4}$$

and the largest positive and negative numbers are usually quoted to specify the static performance of a DAC.

However, several definitions of INL may result depending on how two endpoints are defined. In some DAC architectures the two endpoints are not exactly 0 and V_{ref}. The nonideal reference point causes an offset error, while the nonideal full-scale range gives rise to a gain error. In most DAC applications, these offset and gain errors resulting from the nonideal endpoints do not matter, and the integral linearity can be better defined in a relative measure using a straight line linearity concept rather than the end point linearity in the absolute measure. The straight line can be defined as two endpoints of the actual DAC output voltages or as a theoretical straight line adjusted to best fit the actual DAC output characteristics. The former definition is sometimes called endpoint linearity, while the latter is called best straight line linearity.

10.1.2.2 Monotonicity

The DAC output should increase over its full range as the digital input word to the DAC increases. That is, the negative DNL should be <-1 LSB for a DAC to be monotonic. Monotonicity is critical in most applications, in particular, in digital control applications. The source of nonmonotonicity is an inaccuracy in binary weighting of a DAC. For example, the most significant bit (MSB) has a weight of one half of the full range. If the MSB weight is smaller than the ideal value, the analog output change can be smaller than the ideal step $V_{\text{ref}}/2^N$ when the input digital word changes from $0111\ldots11$ to $1000\ldots00$ at the midpoint of the DAC range. If this decrease in the output is >1 LSB, the DAC becomes nonmonotonic. The similar nonmonotonicity can take place when switching the second or lower MSB bits in binary-weighted multibit DACs.

Monotonicity is inherently guaranteed if an N-bit DAC is made of 2^N elements for thermometer decoding. However, it is impractical to implement high-resolution DACs using 2^N elements because the number of elements grows exponentially as N increases. Therefore, to guarantee monotonicity in practical applications, DACs have been implemented using either a segmented DAC or an

integrator-type DAC. Oversampling interpolative DACs also achieve monotonicity using a pulse-density modulated bitstream converted into analog voltages by a lossy integrator or by a low-pass filter.

Segmented DACs. Applying a two-step conversion concept, a DAC can be made in two levels using coarse and fine DACs. The fine DAC divides one coarse MSB segment into fine LSBs. If one fixed MSB segment is subdivided to generate LSBs, matching among MSB segments creates a nonmonotonicity problem. However, if the next MSB segment is subdivided instead of the fixed segment, the segmented DAC can maintain monotonicity regardless of the MSB matching. This is called ''next-segment approach.'' Unless the next segment approach is used to make a segmented DAC with a total $M+N$ bits, the MSB DAC should have a resolution of $M+N$ bits for monotonicity, while the LSB DAC requires an N-bit resolution. Using the next-segment approach, an MSB DAC made of 2^M identical elements guarantees monotonicity, although INL is still limited by the MSB matching.

To implement a segmented DAC using two resistor-string DACs, voltage buffers are needed to drive the LSB DAC without loading the MSB DAC. Although the resistor-string MSB DAC is monotonic, overall monotonicity is not guaranteed due to the offsets of the voltage buffers. The use of a capacitor-array LSB DAC eliminates a need for voltage buffers. The most widely used segmented DAC is a current-ratioed DAC, whose MSB DAC is made of identical elements for the next-segment approach, but the LSB DAC is a current divider. A binary-weighted current divider can be used as an LSB DAC, as shown in Figure 10.2. For monotonicity, the MSB M-bits are selected by a thermometer code, but one of the MSB current sources corresponding to the next segment of the thermometer code is divided by a current divider for fine LSBs.

Integrator-type DACs. As mentioned, monotonicity is guaranteed only in a thermometer-coded DAC. The thermometer coding of a DAC output can be implemented either by repeating identical DAC elements many times or by repeatedly using the same element. The former requires more hardware, but the latter more time. In the continuous time integrator-type DAC the integrator output is a linear ramp and the time to stop integrating can be controlled by the digital codes. Therefore, monotonicity can be maintained. Similarly, the discrete time integrator can integrate a constant amount of charge repeatedly and the number of integrations can be controlled by the digital codes. The integration approach can give high accuracy, but its disadvantage is its slow speed limiting its applications.

Although it is different in concept, oversampling interpolative DACs modulate the digital code in to a bitstream, and its pulse density represents the DAC output. Due to the incremental nature of the pulse density modulation, oversampling DACs are monotonic. A DAC for the pulse-density modulated

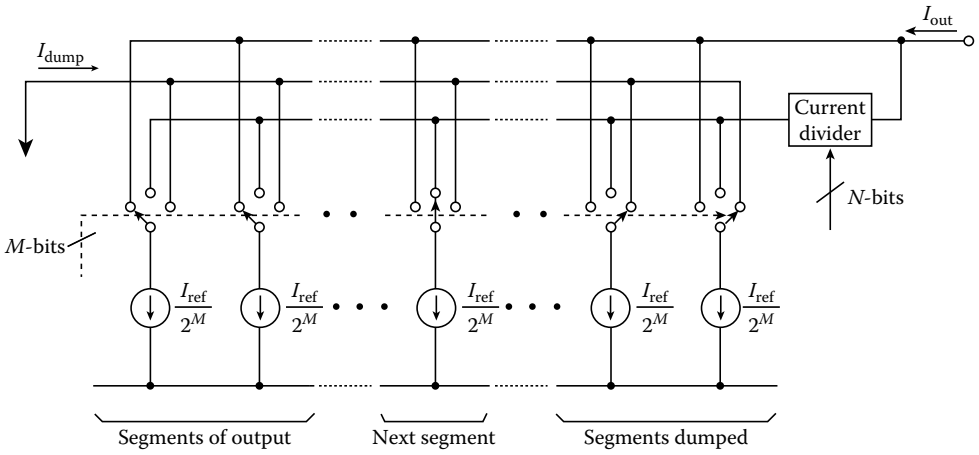


FIGURE 10.2 Segmented DAC for monotonicity.

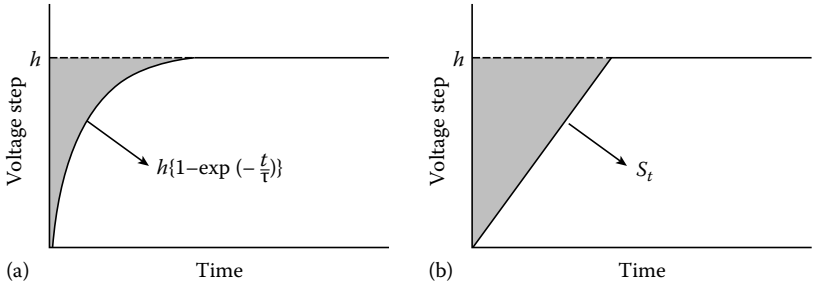


FIGURE 10.3 Errors in step response: (a) settling and (b) slewing.

bitstream is a lossy integrator. The integrator integrates a constant charge if the pulse is high, while it subtracts the same charge if the pulse is low. In principle, it is equivalent to the discrete time integrator DAC, but the output is represented by the average charge on the integrator.

10.1.2.3 Conversion Speed

The output of a DAC is a sampled-and-held step waveform held constant during a word clock period. Any deviation from the ideal step waveform causes an error in the DAC output. High-speed DACs usually have a current output, but even current-output DACs are either terminated with a 50 to 75 Ω low-impedance load or buffered by a wideband transresistance amplifier. Therefore, the speed of a DAC is limited either by the RC time constant of the output node or by the bandwidth of the output buffer amplifier. Figure 10.3 illustrates two step responses of a DAC when it settles with a time constant of τ and when it slews with a slew rate of S, respectively. The transient errors given by the shaded areas of Figure 10.3 are $h\tau$ and $h^2/2S$, respectively. This implies that a single time-constant settling of the former case only generates a linear error in the output, which does not effect the DAC linearity, but the slew-limited settling of the buffer generates a nonlinear error. Even in the single-time constant case (the former), the code-dependant time constant in settling can introduce a nonlinearity error because the settling error is a function of the time constant τ. This is true for a resistor-string DAC, which exhibits a code-dependent settling time because the output resistance of the DAC depends on the digital input.

The slew rate limit of the buffer is a significant source of nonlinearity since the error is proportional to the square of the signal, as shown in Figure 10.3b. The height and the width of the error term change with the input. The worst-case HD when generating a sinusoidal signal with a magnitude V_0 with a limited slew rate of S is [1]

$$\mathrm{HD}_k = 8\frac{\sin^2\frac{\omega T_c}{2}}{\pi k(k^2-4)} \times \frac{V_0}{ST_c}, \quad k = 1, 3, 5, 7, \ldots \tag{10.5}$$

where T_c is the clock period. For a given distortion level, the minimum slew rate is given. Any exponential system with a bandwidth of ω_0 gives rise to signals with the maximum slew rate of $2\omega_0 V_0$. Therefore, by making $2\omega_0 V_0 > S_{\min}$, no slew rate is limited and the DAC system will exhibit no distortion.

10.1.3 Converter Architectures

Many circuit techniques are used to implement DACs, but a few popular techniques used widely today are of the parallel type, in which all bits change simultaneously upon applying an input code word. Serial DACs, on the other hand, produce an analog output only after receiving all digital input data in a sequential form. When DACs are used as stand-alone devices, their output transient behavior limited by

glitch, slew rate, word clock jitter, settling, etc. are of paramount importance, but are used as subblocks of ADCs, DACs need only to settle within a given time interval. In stand-alone DAC applications, the digital input word made of *N*-bits should be synchronously applied to the DAC with a precise timing accuracy. Thus, the input data latches are used to hold the digital input during the conversion. The output analog sample-and-hold (S/H), usually called deglitcher, is often used for the better transient performance of a DAC. The three most popular architectures in integrated circuits are DACs using a resistor string, ratioed current sources, and a capacitor array. The current-ratioed DAC finds the greatest application as a stand-alone DAC, while the resistor-string and capacitor-array DACs are used mainly as ADC subblocks.

10.1.3.1 Resistor-String DACs

The simplest voltage divider is a resistor string. Reference levels can be generated by connecting 2^N identical resistors in series between V_{ref} and 0. Switches to connect the divided reference voltages to the output can be either a 1-out-of-2^N decoder or a binary tree decoder as shown in Figure 10.4 for the 3 bit example. Because it requires a good switch, the stand-alone resistor-string DAC is easier to implement using CMOS. However, the lack of switches does not limit the application of the resistor string as a voltage reference divider subblock for ADCs in other process technologies.

Resistor strings are used widely as reference dividers, an integral part of the flash ADC. All resistor-string DACs are inherently monotonic and exhibit good differential linearity. However, they suffer from a poor integral linearity and also have the drawback that the output resistance depends on the digital input code. This causes a code-dependent settling time when charging the capacitive load of the output bus. The code-dependent settling time has no effect on the reference divider performance as an ADC subblock, but the performance is severely degraded as a stand-alone DAC. This nonuniform settling time problem can be alleviated by adding low-resistance parallel resistors and by compensating the MOS switch overdrive voltages.

In bipolar technology, the most common resistors are thin-film resistors made of tantalum, Ni–Cr, or Cr–SiO, which exhibit very low voltage and temperature coefficients. In CMOS either diffusion or undoped poly resistors are common. Four of the most frequently used resistors are listed in Table 10.1. Conventional trimming or adjustment techniques are impractical to be applied to all 2^N resistor elements. The following four methods are often used to improve the integral linearity of resistor-string DACs.

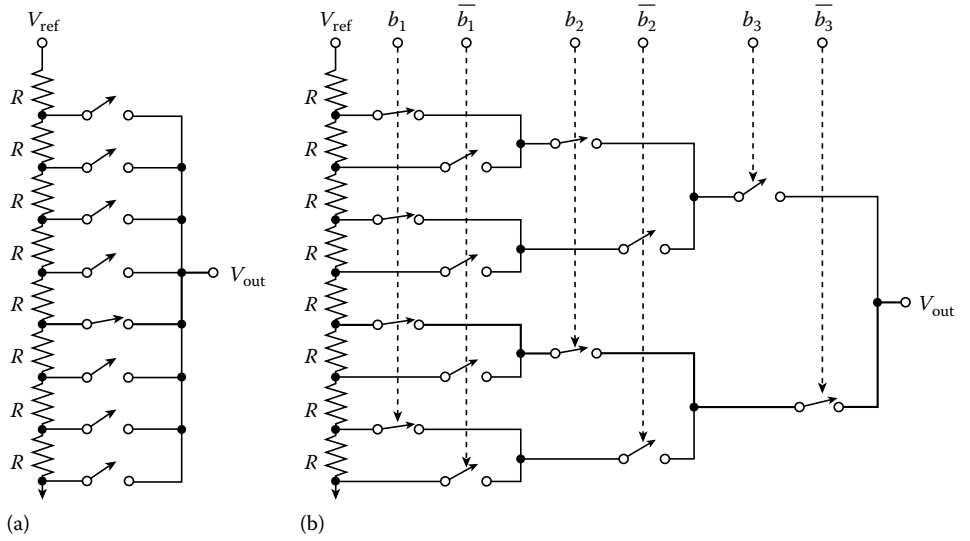


FIGURE 10.4 Resistor-string DAC: (a) with 1-out-of-2^N decoder and (b) with a binary tree decoder.

TABLE 10.1 Resistor in IC Processes

Resistor Type	Sheet R (Ω/sq)	Tolerance (%)	10–20 μm Matching (%)	T.C. (ppm/°C)
Diffusion	100–200	±20	±0.2–0.5	1500
Ion implantation	500–1 k	±5	±0.1	200
Thin film	1 k	±5	±0.1	10–100
Undoped poly	100–500	±20	±0.2	1500

Layout techniques. The use of large geometry devices and/or careful layout is effective in improving the matching marginally. Large geometry devices reduce the random edge effect, and the layout using a common centroid or geometric averaging can reduce the process gradient effect. However, typical matching of resistors in integrated circuits is still limited to an 8–10 bit level due to the mobility and resistor thickness variations. Differential resistor DACs with large feature sizes are reported to exhibit a higher matching accuracy of an 11–12 bit level.

Off-chip adjustment. It is possible to set tap points of a resistor-string to specified voltages by connecting external voltage sources to them, as shown in Figure 10.5a for the 3 bit example. Simply put, the more taps adjusted, the better the integral linearity obtained. An additional benefit of this method is the reduced RC time constant due to the voltage sources at the taps. Instead of using voltage sources, the required voltages can be obtained using parallel trimming, resistors, as shown in Figure 10.5b. However, in addition to external components for trimming, fine adjustments and precision measurement instruments are needed to ensure that voltage levels are correct. Furthermore, due to mismatch in the temperature coefficients between the external components and the on-chip components, retrimming is often required when temperature changes.

Postprocess trimming. The most widely used methods are laser trimming [2], Zener zapping [3], and other electrical trimming using PROM. The trimming method is the same as the parallel resistor trimming shown in Figure 10.5b except for the fact that external trimming resistors are now integrated on the chip. While being trimmed, the resistor string is biased with a constant current. Individual

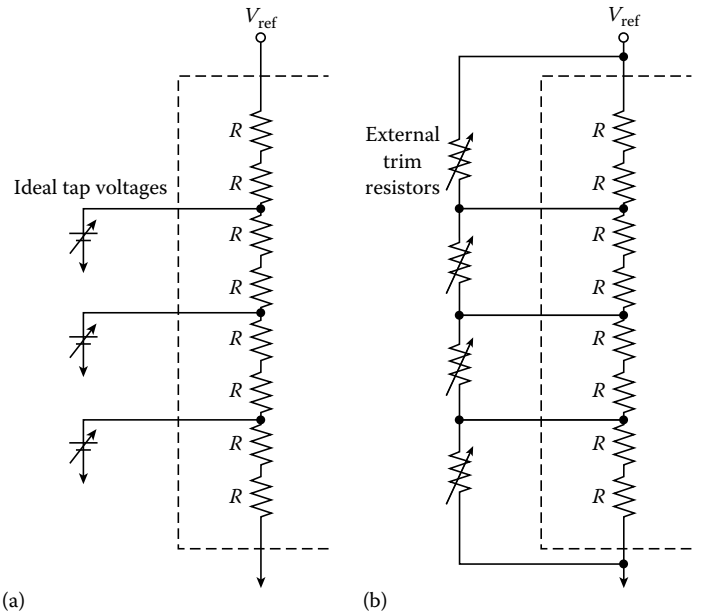


FIGURE 10.5 INL improvements by (a) external voltage sources and (b) parallel resistors.

segments are trimmed to have the same voltage drop. However, during normal conversion, the current source is replaced by a reference voltage source. The focused laser beam for trimming has a finite diameter, and the resistor to be trimmed occupies a large chip area. Both the laser trimming and the Zener zapping processes are irreversible. The long-term stability of trimmed resistors is a major concern, although the electrical and the PROM trimming (if PROM is replaced by EPROM) can be repeated. All trimming methods in this category are time consuming and require precision instruments.

On-chip buffers. The voltage at intermediate taps of the resistor string can be controlled by another resistor string through on-chip unity-gain buffers. This is actually an improved version of the off-chip method. The controlling resistors can be either laser trimmed or electronically controlled by switches. Laser-trimmed controlling resistors have the same problems mentioned earlier. The trimming network can be implemented to electronically control resistor values. In either case buffers with a high open-loop gain, a low output resistance, a large current driving capability, and a wide bandwidth for accurate and fast settling are required.

10.1.3.2 Current-Ratioed DACs

The most popular stand-alone DACs in use today are current-ratioed DACs, of which the two types are a weighted current DAC and an *R*-2*R* DAC.

Binary-weighted current DACs. The weighted current DACs shown in Figure 10.6 are made of an array of switched binary-weighted current sources and the current summing network. In bipolar

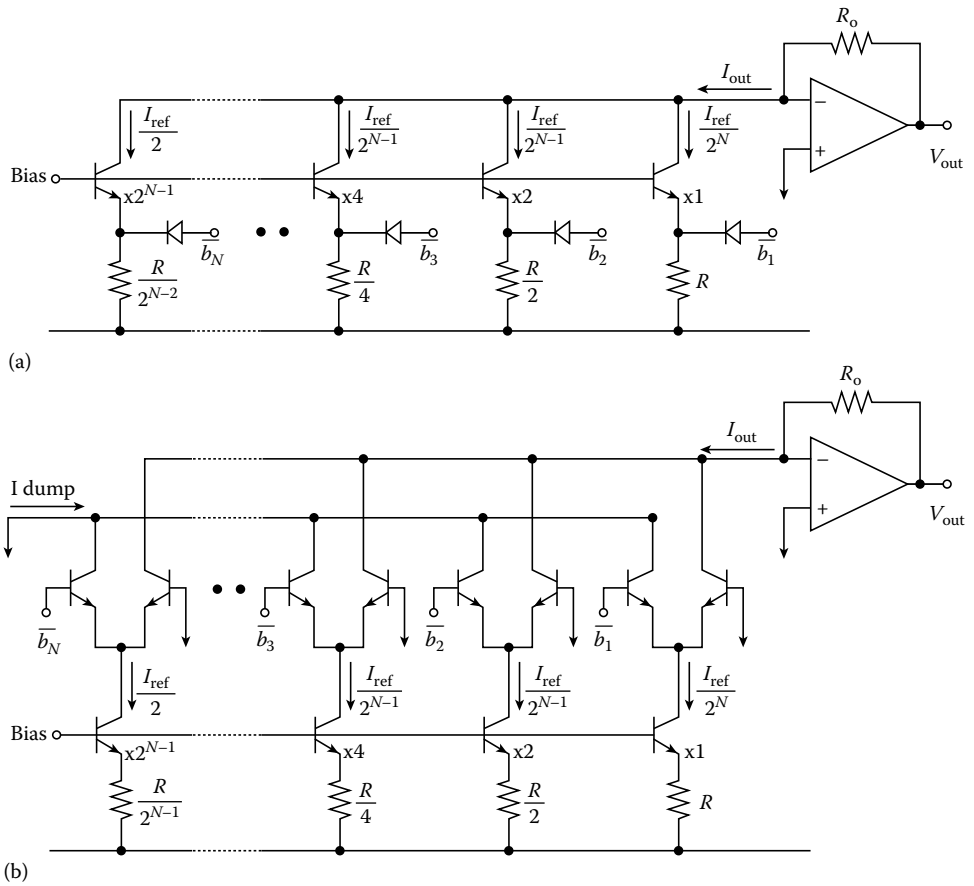


FIGURE 10.6 Binary-weighted current DAC: (a) diode switching and (b) differential pair switching.

technology binary weighting is achieved by ratioed transistors and emitter resistors with binary-related values of R, $R/2$, $R/4$, and so on, while in MOS technology, only ratioed transistors are used. One example is a video random access memory DAC in CMOS, which is made of simple PMOS differential pairs with binary-weighted tail currents. DACs relying on active device matching can achieve an 8 bit level performance with a 0.2%–0.5% matching accuracy using a 10–20 μm device feature size while degeneration with thin-film resistors gives a 10 bit level performance. The current sources are switched on or off by means of switching diodes or emitter-coupled differential pairs (source-coupled pairs in CMOS), as shown in Figure 10.6. The output current summing is done by a wideband transresistance amplifier, but in high-speed DACs, the output current is used directly to drive a resistor load for a maximum speed but with a limited output swing. The weighted current design has the advantage of simplicity and high speed, but it is difficult to implement a high-resolution DAC because a wide range of emitter resistors and transistor sizes is used and very large resistors cause problems with both temperature stability and speed.

***R*-2*R* ladder DACs.** This large resistor ratio problem is alleviated by using a resistor divider known as an *R*-2*R* ladder, as shown in Figure 10.7. The *R*-2*R* network consists of series resistors of value *R* and shunt resistors of value 2*R*. The top of each shunt resistor 2*R* has a single-pole double-throw electronic switch which connects the resistor either to ground or to the output current summing node. The operation of the *R*-2*R* ladder network is based on the binary division of current as it flows down the ladder. At any junction of series resistor *R*, the resistance looking to the right side is 2*R*. Therefore, the input resistance at any junction is *R*, and the current splits into two equal parts at the junction because it sees equal resistances in either direction. The result is binary-weighted currents flowing into each shunt resistor in the ladder. The digitally controlled switches direct the current to either ground or to the summing node. The advantage of the *R*-2*R* ladder method is that only two values of resistors are used, greatly simplifying the task of matching or trimming and temperature tracking. In addition, for high-speed applications relatively low resistor values can be used. Excellent results can be obtained using laser-trimmed thin-film resistor networks. Because the output of the *R*-2*R* DAC is the product of the reference voltage and the digital input word, the *R*-2*R* ladder DAC is often called a multiplying DAC (MDAC).

Both the weighted current DAC and the *R*-2*R* DAC can be used as a current divider to make a sub-DAC. To make a segmented DAC for monotonicity based on the next-segment approach, as discussed earlier, the MSB should be made of thermometer-coded equal currents. Once the MSB is selected, the next segment should be divided further into LSBs as shown in Figure 10.2. INL can be improved by dynamically matching or by self-calibrating the MSB current sources as discussed later.

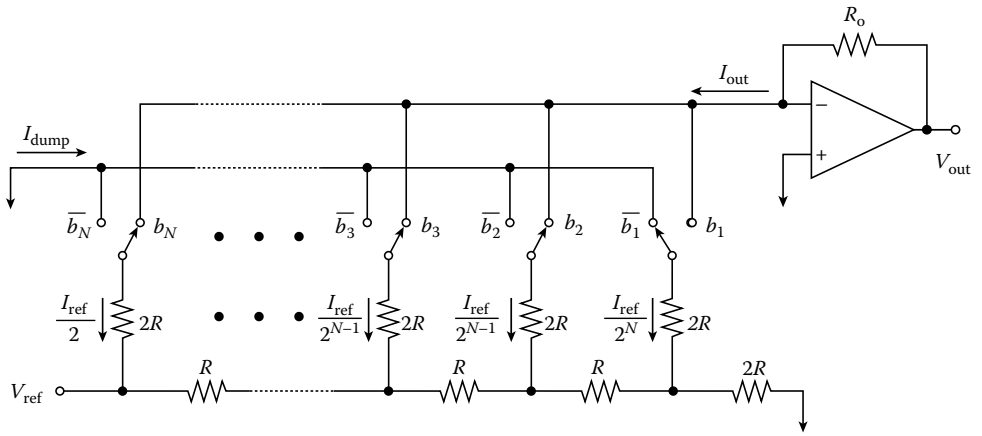


FIGURE 10.7 *R*-2*R* ladder DAC.

Capacitor-array DAC. Capacitors made of double-poly or poly-diffusion in MOS technology are considered one of the most accurate passive components comparable to thin-film resistors in bipolar process, both in the matching accuracy and voltage and temperature coefficients [4]. The only disadvantage in the capacitor-array DAC implementation is the use of a dynamic charge redistribution principle. A switched-capacitor counterpart of the resistor-string DAC is a parallel capacitor array of 2^N unit capacitors (C) with a common top plate. The capacitor-array DAC is not appropriate for stand-alone applications without a feedback amplifier virtually grounding the top plate and an output S/H or deglitcher. The operation of the capacitor-array DAC in Figure 10.8 is based on the thermometer-coded DAC principle, and has a distinct advantage of monotonicity if the system is implemented properly. However, due to the complexity of handling the thermometer-coded capacitor array, a binary-weighted capacitor array is often used, as shown in Figure 10.9, by grouping unit capacitors in binary ratio values. A common centroid layout of the capacitor array is known to give a 10 bit level matching for this application when the unit capacitor size is over 12×12 μm. The matching accuracy of the capacitor in MOS technology depends on the geometry sizes of the capacitor width and length and the dielectric thickness.

As a stand-alone DAC, the top plate of the DAC is precharged either to the offset of the feedback amplifier or to the ground. One smallest capacitor is not necessary for this application. However, as a subblock of an ADC, the total capacitor should be 2^NC, as drawn in Figure 10.9, and the top plate of the

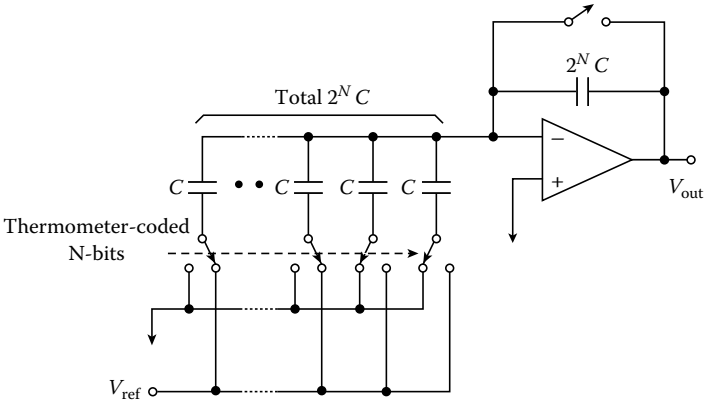


FIGURE 10.8 Thermometer-coded capacitor-array DAC.

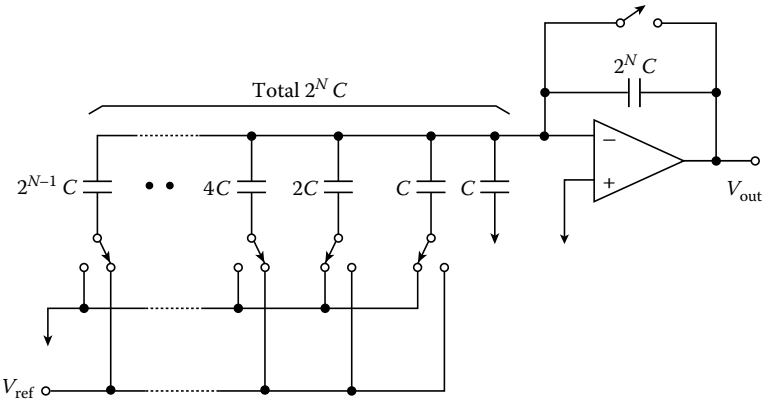


FIGURE 10.9 Binary-weighted capacitor-array DAC.

array is usually connected to the input nodes of comparators or high-gain operational amplifiers, depending on the ADC architectures. As a result, the top plate has a parasitic capacitance, but its effect on the DAC performance is negligible. The capacitor-array DAC requires two-phase nonoverlapping clocks for proper operation. Initially, all capacitors should be charged to ground. After initialization, depending on the digital input, the bottom plates are connected either to V_{ref} or to ground. Consider the case in which the top plate is floating without the feedback amplifier. If the charge at the top plate finishes its redistribution the top plate voltage neglecting the top plate parasitic effect becomes

$$V_0 = \sum_{i=1}^{N} \frac{b_k}{2^i} V_{\text{ref}} \tag{10.6}$$

where $b_N\, b_{N-1}, \ldots, b_2\, b_1$ is the input binary word. For example, switching the MSB capacitor bottom to V_{ref} changes the output voltage by

$$\frac{2^{N-1}C}{\sum_{i=1}^{N} C_i} V_{\text{ref}} = \frac{V_{\text{ref}}}{2} \tag{10.7}$$

where the capacitor C_i, for the ith bit is nominally scaled to $2^{i-1}C$. Therefore, the nonlinearity at the midpoint of the full range is limited by the ratio mismatch of the half sum of the capacitor array to the total sum of the array. Similarly, the nonlinearity at one fourth of the range is limited by the ratio of one fourth of the capacitor array to the total array, and so on.

One important application of the capacitor-array DAC is as a reference DAC for ADCs. As in the case of the *R*-2*R* MDAC, the capacitor-array DAC can be used as an MDAC to amplify residue voltages for multistep ADCs. As shown in Figure 10.9, if the input is sampled on the bottom plates of capacitors instead of the ground, the voltage amplified is the amplified input voltage minus the DAC output. By varying the feedback capacitor size, the MDAC can be used as an interstage residue amplifier in multistep pipelined ADCs. For example, if the feedback capacitor is C and the digital input is the coarse N-bit decision of the sampled analog voltage, the amplifier output is a residue voltage amplified by 2^N for the subsequent LSB conversion.

10.1.3.3 $R + C$ or $C + R$ Combination DACs

Both resistor-string and capacitor-array DACs need 2^N unit elements for N-bits, and the number grows exponentially. Splitting arrays into two, one for MSBs and the other for the LSBs, requires a buffer amplifier to interface between two arrays. Although a floating capacitor connects two capacitors arrays, the parasitic capacitance of the floating node is not well controlled. A more logical combination for high-resolution DAC is between resistor and capacitor DACs. This combination does not require any coupling capacitors or interface buffer amplifiers.

In the $R + C$ combination, the MSB is set by the resistor string, and next segment of the resistor-string DAC supplies the reference voltage of the LSB capacitor DAC, as shown in Figure 10.10. When the top plate is initialized, all capacitor bottom plates are connected to the lower voltage of the next segment of the resistor-string DAC. During the next clock phase, the bottom plates of capacitors are selectively connected to the higher voltage of the segment if the digital bit is ONE, but stays switched to the lower voltage if ZERO. This segmented DAC approach gives an inherent monotonicity as far as the LSB DAC is monotonic within its resolution. Although INL is poor, the fully differential implementation of this architecture benefits from the lack of the even-order nonlinearity, thereby achieving improved INL. On the other hand, in the $C + R$ combination shown in Figure 10.11, the operation of the capacitor DAC is the same. The MSB side reference voltage is fixed, but the reference voltage of the smallest capacitor is supplied by the LSB resistor-string DAC. This approach exhibits nonmonotonicity due to the capacitor

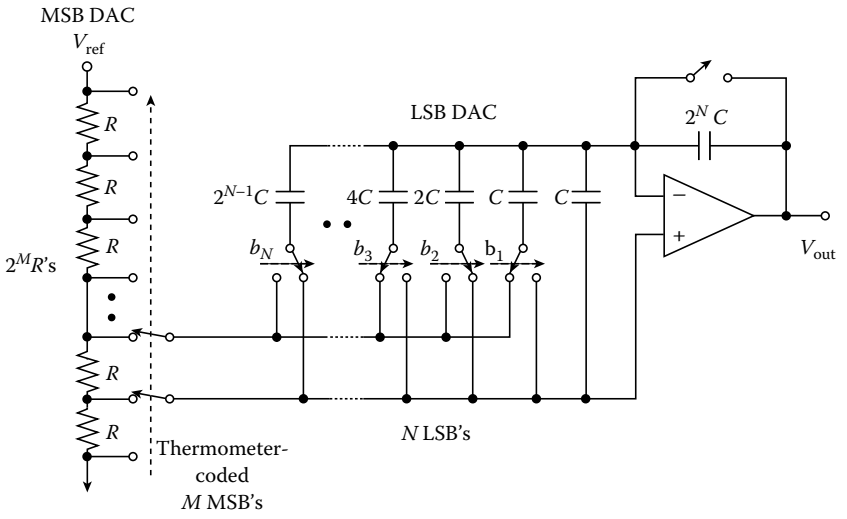


FIGURE 10.10 $R + C$ combination DAC.

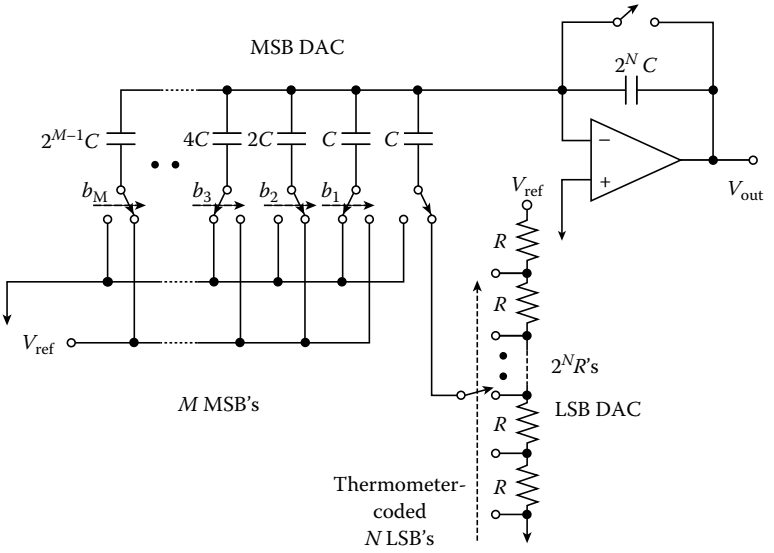


FIGURE 10.11 $C + R$ combination DAC.

DAC matching. Both combination DACs are seldom used as stand-alone DACs due to their limited speed, but are used frequently as subblocks of high-resolution ADCs.

10.1.4 Techniques for High-Resolution DACs

Most DACs are made of passive or active components such as resistors, capacitors, or current sources, and their linearity relies on the matching accuracy of those components. Among frequently used DAC components, diffused resistors and transistors are in general known to exhibit an 8 bit level matching while thin-film resistors and capacitors are matched to a 10 bit level. Trimming or electronic calibration is needed in order to obtain a higher linearity than what is achievable with bare component matching.

The traditional solutions to this have been the wafer-level trimming methods such as laser trimming and Zener zapping. Although many other promising trimming or matching techniques such as polysilicon fuse trimming, electrical trimming techniques using PROM, and large device matching by geometrical averaging have been proposed, conventional factory-set trimming or matching techniques give no flexibility of retrimming. How successfully these techniques can be applied to large-volume production of high-resolution DACs and how the factory-trimmed components will perform over the long term are still in question.

The future trend is toward more sophisticated and intelligent electronic solutions that overcome and complement some of the limitations of conventional trimming techniques. The methods recently developed are dynamic circuit techniques [5] for component matching, switched-capacitor integration [6], electronic calibration [7] of a DAC nonlinearity, and oversampling interpolation techniques [8] which trade speed with resolution. In particular, the oversampling interpolative DACs are used widely in stand-alone applications such as digital audio playback systems or digital communications due to their inherent monotonicity.

10.1.4.1 Dynamic Matching Techniques

In general, a dynamic element matching to improve the accuracy of the binary ratio is a time-averaging process. For simplicity, consider a simple voltage or current divide-by-two element, as shown in Figure 10.12. Due to mismatches in the ratio of resistors, transistors, and capacitors, the divided voltage or current is not exactly $V_{ref}/2$ or $I_{ref}/2$, but their sum is V_{ref} or I_{ref}. The dynamic matching concept is to multiplex these two outputs with complementary errors of Δ and $-\Delta$ so that the errors Δ and $-\Delta$ can be averaged out over time while the average value of $V_{ref}/2$ or $I_{ref}/2$ remains. It is in effect equivalent to the suppressed carrier balanced modulation of the error component Δ. The high-frequency energy can be filtered out using a post low-pass filter. This technique relies on the accurate timing of the duty cycle. Any duty cycle error or timing jitter results in inaccurate matching. The residual matching inaccuracy becomes a second-order error proportional to the product of the original mismatch and the timing error.

The application of dynamic element matching to the binary-weighted current DAC is a straightforward switching of two complementary currents. Its application to the binary voltage divider using two identical resistors or capacitors requires exchanging of resistors or capacitors. This can be achieved easily by reversing the polarity of the reference voltage for the divide-by-two case. However, in the general case of N element matching the current division is inherently simpler in implementation than the voltage division. In general, to match the N independent elements, a switching network with N inputs and N outputs is required. The function of the switching network is to connect any input out of N inputs to

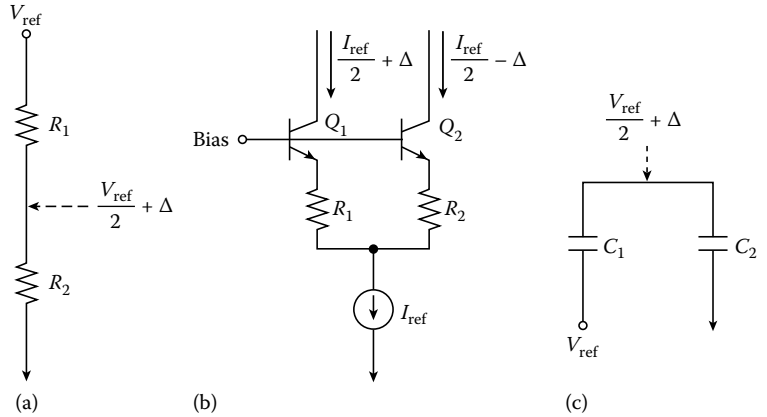


FIGURE 10.12 Divide-by-two elements: (a) resistor, (b) current, and (c) capacitor.

one output with an average duty cycle of $1/N$. The simplest one is a barrel shifter rotating the I–O connections in a predetermined manner [5]. This barrel shifter generates a low-frequency modulated error when N becomes larger because the same pattern repeats every N clocks. A more sophisticated randomizer with the same average duty can distribute the mismatch error over the wider frequency range. The latter technique finds applications as a multibit DAC in the multibit noise-shaping sigma-delta data converter, whose linearity relies on the multibit DAC.

Voltage or current sampling. The voltage or current sampling concept is an electronic alternative to direct mechanical trimming. To sample voltage or current using a voltage or current sampler is equivalent to trimming individual voltage or current sources. The voltage sampler is usually called a S/H circuit, while the current sampler is called a current copier. The voltage is usually sampled on the input capacitor of a buffer amplifier and the current is usually sampled on the input capacitor of a transconductance amplifier such as an MOS transistor gate. Therefore, both voltage and current sampling techniques are ultimately limited by their sampling accuracy.

The idea behind the voltage or current sampling DAC is to use one voltage or current element repeatedly. One example of the voltage sampling DAC is a discrete-time integrator-type DAC with many S/H amplifiers for sampling output voltages. The integrator integrates a constant charge repeatedly, and its output is sampled on a new S/H amplifier every time the integrator finishes an integration as shown in Figure 10.13a. This is equivalent to generating equally spaced reference voltages by stacking identical unit voltages [6]. The fundamental problem associated with this sampling voltage DAC approach is the accumulation of the sampling error and noise in generating larger voltages. Similarly, the current sampling DAC can sample a constant current on current sources made of MOS transistors, as shown in Figure 10.13b [7]. Because one reference current is copied on other identical current samplers, the matching accuracy can be maintained as far as the sampling errors are kept constant. It is not practical to make a high-resolution DAC using voltage or current sampling alone. Therefore, this approach is limited to generating MSB DACs for the segmented DAC or for the subranging ADCs.

10.1.4.2 Electronic Calibration Techniques

Electronic calibration is a general term to describe various circuit techniques, which usually predistort the DAC transfer characteristic so that the DAC linearity can be improved. One such technique is a

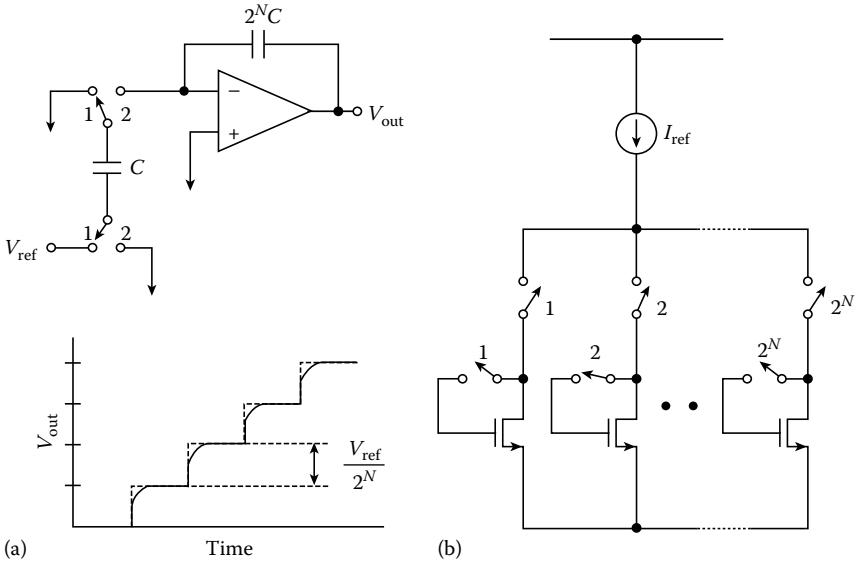


FIGURE 10.13 Voltage and current sampling concepts: (a) integrator and (b) current copier.

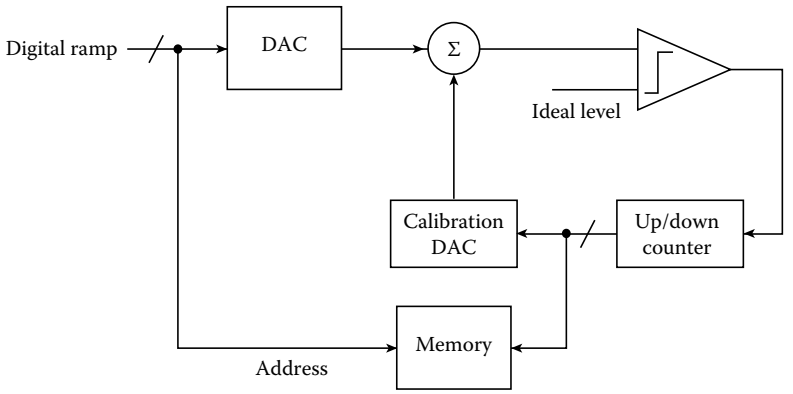


FIGURE 10.14 Code mapping with a calibration DAC.

straightforward code mapping, and the other is a self-calibration. The code-mapping calibration is a very limited technique only for the factory because it requires a precision measurement setup and a large digital memory. The self-calibration is to incorporate all the calibration mechanisms and hardware on the DAC as a built-in function so that users can recalibrate whenever calibrations are needed. The self-calibration is based on an assumption that the segmented DAC linearity is limited by the MSB DAC so that only errors of MSBs may be measured, stored in memory, and recalled during normal operation. There are two ways of measuring the MSB errors. In one method individual-bit nonlinearities, usually appearing as component mismatch errors, are measured digitally [9], and a total error, which is called a code error, is computed from individual-bit errors depending on the output code during normal conversion. On the other hand, the other method measures and stores digital code errors directly and eliminates the digital code-error computation during normal operation [10]. The former requires less digital memory during normal conversion while the latter requires fewer digital computations.

Direct code mapping. The simplified code mapping of a DAC can be done with a calibration DAC, digital memory, and a precision instrument to measure the DAC output, as shown in Figure 10.14. The idea is to measure the DAC error using a calibration DAC so that the DAC output corrected by the calibration DAC can produce an ideal DAC output. The input code of the calibration DAC is stored as a code error in digital memory addressed by the DAC input code. This code error is recalled to predistort the DAC output during normal operation. This technique needs a 2^N memory with a word length corresponding to the number of bits of the calibration DAC. This method can correct any kinds of DAC nonlinearities as long as the calibration DAC has an output range wide enough to cover the whole range of nonlinearity. However, the same method can be implemented without the use of a calibration DAC if the main DAC is monotonic with extra bits of resolution. In this case, the calibration is a simple code mapping, selecting correct input digital codes for correct DAC output voltages among redundant input digital codes.

Self-calibration for individual capacitor errors. The idea of measuring the individual bit errors using a calibration DAC is to quantize the difference Δ in the divide-by-two elements in Figure 10.12 because the ideal divide ration is 1/2. For example, the MSB should be half of the whole range of a DAC, the second MSB is half of the MSB, and so on. Unless buffer amplifiers are used, the ideal calibration DACs for *R* and *C* DACs are *C* and *R* DACs, respectively. The ratio error measurement cycle of 2 bit *C* DAC is illustrated in Figure 10.15. Errors can be quantized using a successive approximation method, but the up/down converter is shown here for simplicity. Initially, the top plate is charged to the comparator offset and the bottom plates of C_1 and C_2 sample 0 and V_{ref}. At the same time, the bottom plate of C_C samples $V_{\text{ref}}/2$ and the up/down counter is reset to make $V_{\text{cal}} = 0$. In the next clock period the charge is redistributed by swapping 0 and V_{ref} on the bottom plates of C_1 and C_2. Then the top plate residual error V_x will be from the charge conservation

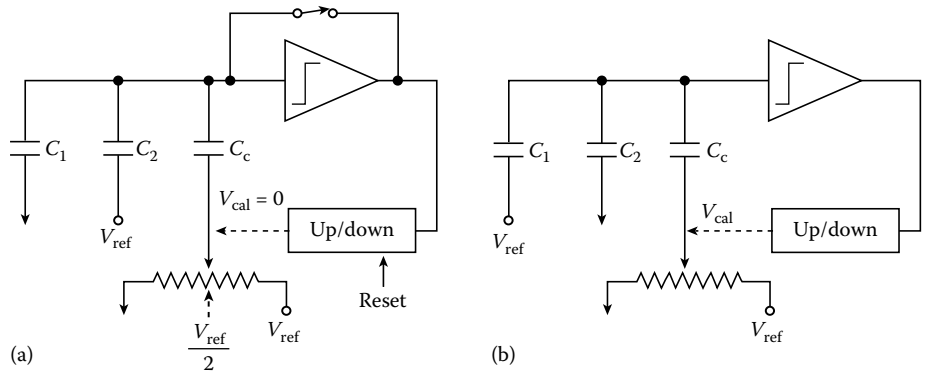


FIGURE 10.15 Capacitor ratio error measurement cycles: (a) initialization and (b) error quantization.

$$V_x = \frac{C_1}{C_1 + C_2 + C_C} V_{ref} - \frac{C_2}{C_1 + C_2 + C_C} V_{ref} = \frac{\Delta C}{2C + C_C} V_{ref} \tag{10.8}$$

If $V_{cal} = 0$, $C_1 = C + \Delta C/2$, and $C_2 = C - \Delta C/2$. This top plate residual voltage can be nulled out by changing the calibration DAC voltage V_{cal}. The measured calibration voltage is approximately

$$V_{cal} = -\frac{\Delta C}{C_C} V_{ref} \tag{10.9}$$

As the actual error V_x is half of the measured value when the V_{ref} is applied to C_1, the actual calibration DAC voltage to be subtracted during normal operation becomes $V_{cal}/2$. Similarly, the multibit calibration can start from the MSB measurement and move down to the LSB side [9].

The extension of this calibration technique to current DACs is straightforward. For example, two identical unipolar currents, I_1 and I_2, can be compared using a voltage comparator and a calibration DAC, as shown in Figure 10.16. After I_1 is switched in, the calibration DAC finds an equilibrium as a null. Then the difference can be measured by interchanging I_1 and I_2 and finding a new equilibrium. Therefore, the current difference error is obtained as

$$I_{cal} = I_2 - I_1 \tag{10.10}$$

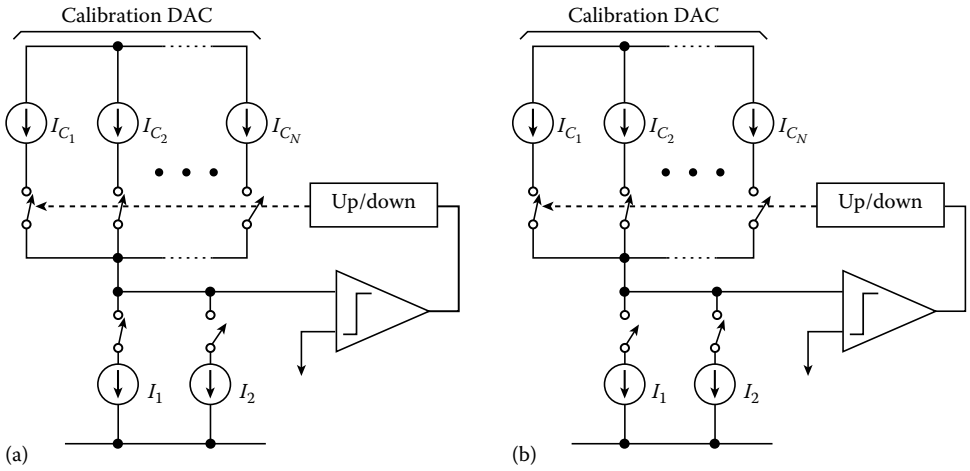


FIGURE 10.16 Current difference measurement cycles: (a) initialization and (b) error quantization.

During normal operation, half of this value should be added to the DAC output using the same calibration DAC every time the current I_1 is switched to the DAC output. Similarly, the same amount is subtracted if I_2 is switched to the output.

Code-error calibration. The code-error calibration is based on the simple fact that the thermometer-coded MSBs of a DAC are made of segments of equal magnitude [10]. Any nonuniform segment will contribute to the overall nonlinearity of a DAC. The segment error between two adjacent input codes is measured by comparing the segment with the ideal segment. Starting from the reference point, 0 or $V_{ref}/2$, the same procedure is repeated until all the segment errors are measured. Therefore, the current code error, Error(j), is obtained by adding the current segment error to the accumulated sum of all the previous segment errors:

$$\text{Error}(j) = \sum_{k=1}^{j} \text{Seg}(k) \tag{10.11}$$

where Seg(k) is the kth segment error from the reference point. These measured code errors are stored in memory addressed by digital codes so that they can be subtracted from uncalibrated raw digital outputs during normal conversion.

The segment error measurement of a current-ratioed DAC thermometer-coded MSBs is similar to the current difference measurement in Figure 10.16. The only difference in measurement is the use of the reference segment current in place of one of the two currents to be compared. That is, each MSB current source is compared to the reference segment. For the capacitive DAC, the kth segment error can be measured in two cycles. After the output of the DAC is initialized to have a negative ideal segment voltage with the input digital code corresponding to $k-1$, the input code is increased by 1 as shown in Figure 10.17. Applying digital codes to the capacitor-array DAC means connecting the bottom plates to either V_{ref} or ground depending on the corresponding digital bits. Then the kth segment error is generated at the output and can be measured digitally using subsequent ADC stages or using a calibration DAC as shown in Figure 10.15.

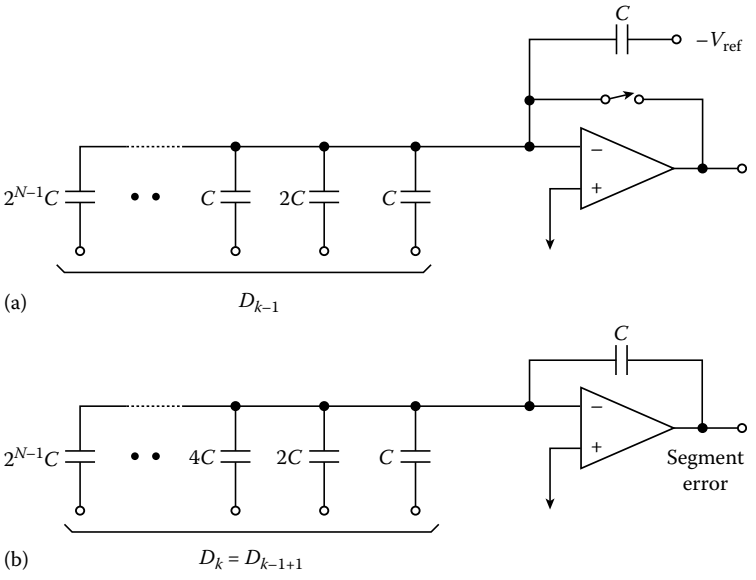


FIGURE 10.17 Code-error measurement cycle: (a) initialization and (b) error quantization.

Digital truncation errors. All calibration methods need extra bits of resolution in the error measurements because digital truncation errors are accumulated during code-error computations. For example, if the truncation errors are random, the additions of n digital numbers will increase the standard deviation of the added number by $n^{1/2}$. This accumulated truncation error affects both DNL and INL of the converter self-calibrated using measured errors of individual bits. On the other hand, if calibrated using measured segment errors, the DNL of the converter is always guaranteed to be within $\pm\frac{1}{2}$ LSB of a target resolution because all segment errors are measured with one extra bit of resolution, but the INL will still be affected by the digital truncation because code errors are obtained by accumulating segment errors. The effect of the digital truncation errors due to n repeated digital additions on the INL can be modeled using uncorrelated and independent random variables, and the standard deviation of INL is calculated in LSB units as

$$\sigma_{\text{INL}} = \sqrt{\frac{(n-i)(i-1)}{12(n-1)}}(\text{LSB}) \quad \text{for } i = 1, 2, \ldots, n \tag{10.12}$$

For example, when $n = 16$, the maximum standard deviation of the INL at the midpoint is about 0.56 LSB.

10.1.4.3 Interpolative Oversampling Techniques

Ordinary DACs generate a discrete output level for every digital word applied to their input, and it is difficult to generate a large number of distinct output levels for long words. The oversampling interpolative DAC achieves fine resolution by covering the signal range with a few widely spaced levels and interpolating values between them. By rapidly oscillating between coarse output levels, the average output corresponding to the applied digital code can be generated with reduced noise in the signal band [8]. The general architecture of the interpolative oversampling DAC is shown in Figure 10.18. A digital filter interpolates sample values of the input signal in order to raise the word rate to a frequency well above the Nyquist rate. The core of the technique is a digital truncator to truncate the input words to shorter output words. These shorter words are then converted into analog form at the high sample rate so that the truncation noise in the signal band may be satisfactorily low. The sampling rate upconversion for this is usually done in stages using two upsampling digital filters. The first filter, usually a two to four times oversampling FIR, is to shape the signal band for sampling rate upconversion and to equalize the passband droop resulting from the second SINC filter for higher-rate oversampling.

A noise-shaping sigma-delta-sigma modulator can be built in digital form to make a digital truncator as shown in Figure 10.19. Using a linearized model, the z-domain transfer function of the modulator is

$$Y(z) = \frac{\alpha H(z)}{1 + \alpha H(z)} X(z) + \frac{1}{1 + \alpha H(z)} Q(z) \tag{10.13}$$

where

$Q(z)$ is the quantization noise
α is the quantizer gain

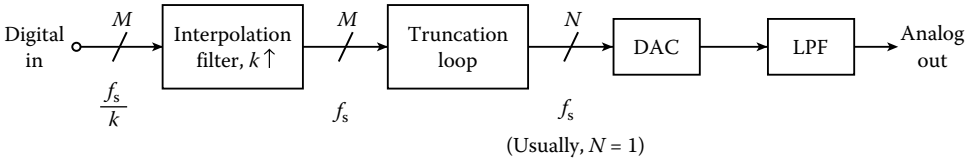


FIGURE 10.18 Interpolative oversampling DAC.

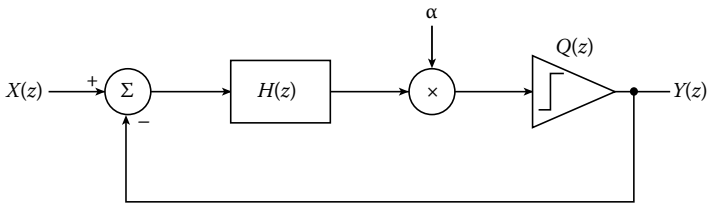


FIGURE 10.19 Delta-sigma modulation as a digital truncator.

The loop filter $H(z)$ can be chosen so that the quantization noise may be high-pass filtered while the input signal is low-pass filtered. For the first-order modulator, the loop filtered is just an integrator with a transfer function of

$$H(z) = \frac{z^{-1}}{1 - z^{-1}} \tag{10.14}$$

while for the second-order modulator, the transfer function is

$$H(z) = \frac{z^{-1}(2 - z^{-1})}{(1 - z^{-1})^2} \tag{10.15}$$

However, the standard second-order modulator is implemented, as shown in Figure. 10.20, using a double integration loop. In general, first-order designs tend to produce correlated idling patterns. Second-order designs are vastly superior to first-order designs both in terms of the required OSR to achieve a particular SNR as well as in the improved randomness of the idling patterns. However, even the second-order loop is not entirely free of correlated fixed patterns in the presence of small DC inputs. The second-order loop needs dithering to reduce fixed pattern noises, but loops of a higher order than third do not exhibit fixed pattern noises.

Stability. The quantizer gain α plays an important role in keeping the modulator stable. Considering α, the transfer function of the second-order loop shown in Figure 10.20 becomes

$$Y(z) = \frac{\alpha z^{-2} X(z) + (1 - z^{-1})^2 Q(z)}{1 - 2(1 - \alpha)z^{-1} + (1 - \alpha)z^{-2}} \tag{10.16}$$

The root locus of the transfer function in the z-domain is shown in Figure 10.21. As shown, the second-order loop becomes unstable for $\alpha > 4/3$ because one pole moves out of the unit circle. This in turn implies that the signal at the input of the quantizer becomes too large. Most delta-sigma modulators become unstable if the signal to the quantizer exceeds a certain limit. Higher-order modulators tend to be

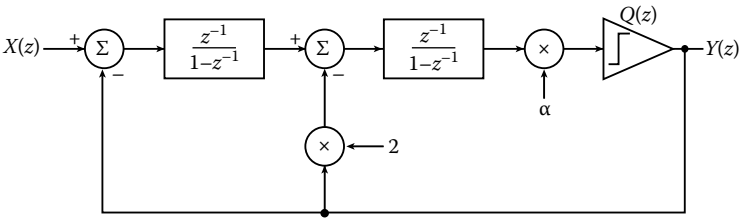


FIGURE 10.20 Second-order 1 bit modulator.

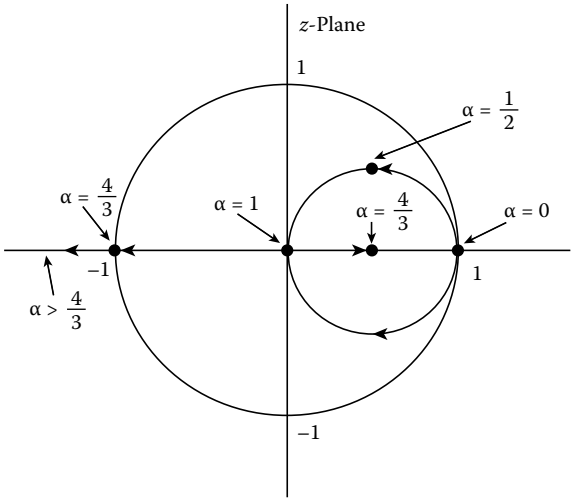


FIGURE 10.21 Root locus of the second-order loop transfer function.

overloaded easily at higher quantizer gain than first-or second-order modulators. Therefore, for the stability reason, the integrator outputs of the loop filter are clamped so that the signal at the input of the quantizer is limited for linear operation. Digital truncators of a higher order than second are feasible in digital circuits because signal levels can be easily detected and controlled. The straightforward third order or higher loop using multiple loops is unstable, but higher order modulators can be built using either the cascaded MASH [11] or the single-bit higher order [12] architecture.

Dynamic range. In general, for the Nth order loop, the noise falls by $6N + 3$ dB for every doubling of the sampling rate, providing $N + 0.5$ extra bits of resolution. Because the advantage of oversampling begins to appear when the OSR is >2, a practically achievable dynamic range by oversampling is approximately

$$\mathrm{DR} > (6N + 3)(v)\mathrm{dB} \tag{10.17}$$

where M is the OSR. For example, a second-order loop with 256 times oversampling can give a dynamic range of >105 dB, but the same dynamic range can be obtained using a third-order loop with only 64 times oversampling. The dynamic range is not a limit in the digital modulator. In practice, the dynamic range is limited in the rear-end analog DAC and postfilter.

One-bit or multibit DAC. The rear end of the interpolative oversampling DAC is an analog DAC. Because the processing in the interpolation filter and truncator is digital instead of analog, achieving precision is easy. Therefore, the oversampling DAC owes its performance to the rear-end analog part because the conversion of the truncated digital words into analog form takes place in the rear-end DAC. The 1 bit quantizer can be easily overloaded and needs clamping to be stable, while multibit quantizers are more stable due to their small quantization errors. However, the multibit system is limited by the accuracy of the multibit DAC. Although the analog techniques such as dynamic matching or self-calibration can improve the performance of the multibit DAC, the 1 bit DAC is simpler to implement and its performance is not limited by component matching. It is true that a continuous time filter can convert the 1 bit digital bitsteam into an analog waveform, but it is difficult to construct an ideal undistorted digital waveform without clock jitter. However, if the bitstream is converted into a charge packet, a high linearity is guaranteed due to the uniformity of the charge packets.

A typical differential 1 bit switched-capacitor DAC with one-pole roll-off can be built as shown in Figure 10.22 using two-phase nonoverlapping clocks 1 and 2. There are many advantages in a fully

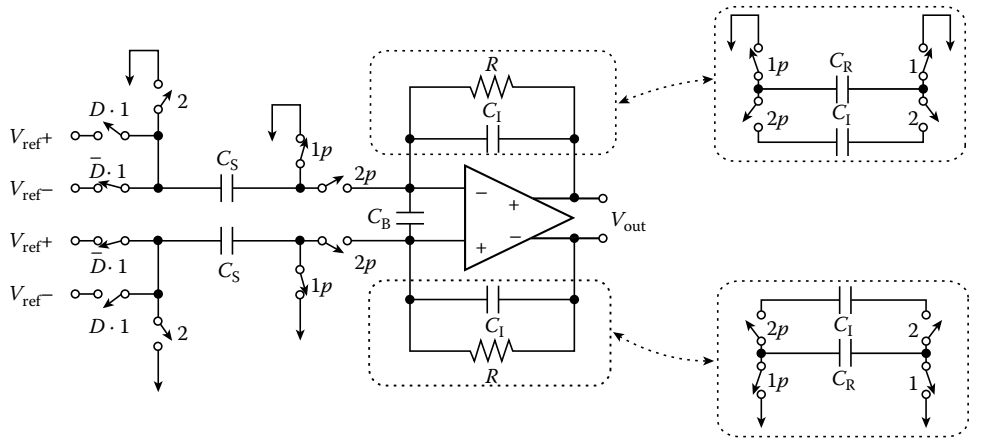


FIGURE 10.22 Switched-capacitor 1 bit DAC/filter.

differential implementation. The dynamic range increases by 3 dB because the signal is doubled (6 dB) but the noise gains by 3 dB. It also rejects most noisy coupling through power supplies or through the substrate as a common-mode signal. Furthermore, the linearity is improved because the even-order nonlinearity components of the capacitors and the op-amp are canceled. In the implementation of Figure 10.22, a resistor as a loss element can be replaced by a capacitor switched in and out at f_c as illustrated. The bandwidths of these filters in both cases are set by $1/RC_I$ and f_cC_R/C_I, respectively, Also, the filter DC gains are defined by Rf_c C_S and C_S/C_R, respectively. The digital bitstream is converted into a charge packet by sampling the reference voltage on the bottom plates of the sampling capacitors (C_S). If the digital bit is ZERO, $-V_{ref}$ is sampled during the clock phase 1 and the charge on C_S is dumped on the lossy integrator during phase 2. On the other hand, if the digital bit is ONE, V_{ref} is sampled instead. To reduce the input-dependent switch-feedthrough component, the switches connected to the top plates should be turned off slightly earlier than the bottom plate switches using $1p$ and $2p$. Alternatively, a slightly different 1 bit DAC is possible by sampling a constant reference voltage by inverting the polarity of the integration depending on the digital bit as shown in Figure 10.23.

The op-amp for this application should have a high DC gain and a fast slew rate. The op-amp DC gain requirement is a little alleviated considering the linear open-loop transfer characteristic of most op-amps

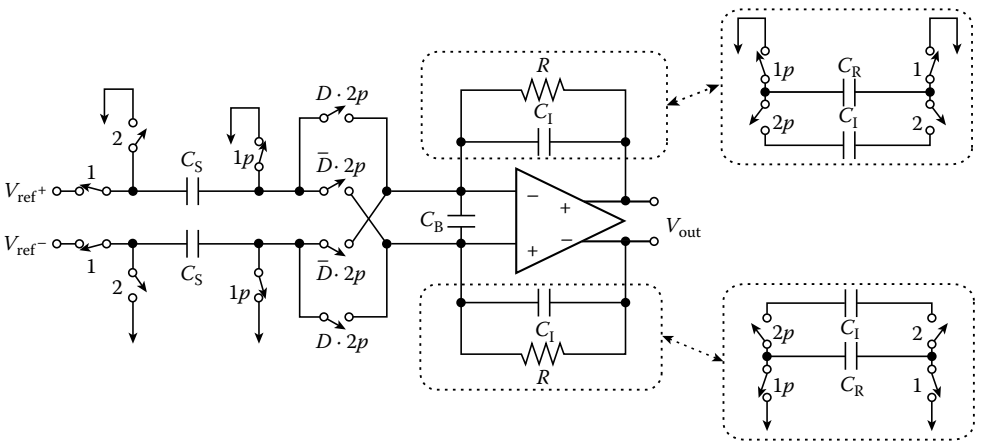


FIGURE 10.23 Alternative 1 bit DAC sampling constant V_{ref}.

within a limited swing range. As discussed earlier, the slew-limited settling generates an error proportional to the square of the magnitude. Therefore, a nonslewing op-amp such as a class AB input op-amp performs better for this application. The op-amp starts to slew when a larger voltage than its linear input range is applied. When the charge packet of the sampled reference voltage is dumped onto the input summing node, it causes a voltage step on the summing node. The bypass capacitor C_B between two summing nodes helps to reduce this voltage step to prevent the op-amp from slewing. The large the C_B, the smaller the voltage step. However, a too-large C_B will narrowband the feedback amplifier, and the settling will take longer as a result.

Postfiltering requirement. Although the one-pole roll-off will substantially attenuate high-frequency components around f_c, the 1 bit DAC should be followed by a continuous time postfilter so that the charge packets can be smoothed out. Unlike the delta-sigma modulator which filters out the out-of-band-shaped noise using digital filters, the demodulator output noise can be filtered only by analog filters. Because the shaped noise is out-of-band, it does not affect the inband performance directly, but the large out-of-band high-frequency noise tends to generate inband intermodulation components and limit the dynamic range of the system. Therefore, the shaped high-frequency noise needs to be filtered with a low-pass filter one order higher than the order of the modulator. It is challenging to meet this postfiltering requirement with analog filtering techniques. Analog filters for this application are often implemented in continuous time using a cascade of Sallen–Key filters made of emitter follower unity-gain buffers, but both switched-capacitor and continuous time filtering techniques have improved significantly to be applied to this application. The other possibility is the hybrid implementation of an FIR filter using digital delays and an analog current swimming network. Because the output is a bitstream, current sources weighted using coefficients of an FIR filter are switched to the current summer depending on the digital bit to make a low-pass FIR filter.

10.1.5 Sources of Conversion Errors

10.1.5.1 Glitch

The basic function of the DAC is the conversion of digital numbers into analog waveforms. A distortion-free DAC creates instantaneously an output voltage that is proportional to the input digital number. In reality, DACs cannot achieve this impossible goal. If the input digital number changes from one value to a different one, the DAC output voltage always reaches a new value sometime later. For DACs, the shape of the transient response is a function governed in large part by two mechanisms, glitch and slew rate limit. The ideal transient response of a DAC to a step is a single-time constant exponential function, which only generates an error growing linearly with the input signal, as explained in Figure 10.3. Any other transient responses give rise to errors that have no bearing on the input signal. The glitch impulse is described in terms of a picovolts times seconds or equivalent unit.

Glitches are caused by small time differences between some current sources turning off and others turning on. Take, for example, the major code transition at half scale from 011 . . . 11 to 100 . . . 00. Here, the MSB current source turns on while all other current sources turn off. The small difference in switching times results in a narrow half-scale glitch, as shown in Figure 10.24. Such a glitch, for example, can produce distorted characters in CRT display applications. To alleviate both glitch and slew-rate problems related to transients, a DAC is followed by a S/H amplifier, usually called a deglitcher. The deglitcher stays in the hold mode while the DAC changes its output value. After the switching transients have settled, the deglitcher is changed to the sampling mode. By making the hold time suitably long, the output of the deglitcher can be made independent of the DAC transient response. Thus, the distortion during transients can be circumvented by using a fast S/H amplifier. However, the slew rate of the deglitcher is on the same order as that of the DAC, and the transient distortion will still be present, now as an artifact of the deglitcher.

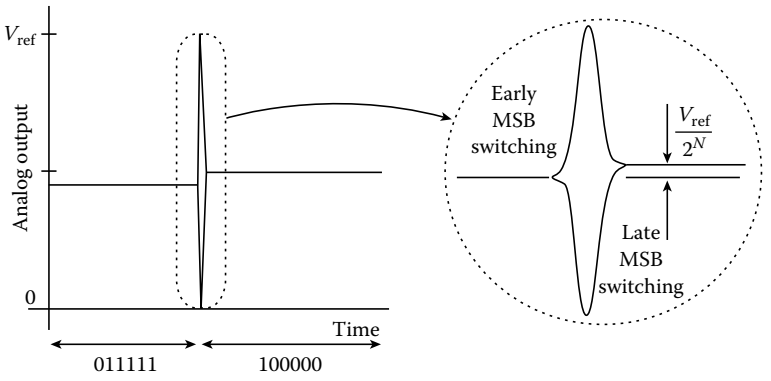


FIGURE 10.24 Glitch impulse at a major carry.

10.1.5.2 Timing Error-Word Clock Jitter

Although a DAC is ideally linear, it needs precise timing to correctly reproduce an analog output signal. If the samples do not generate an analog waveform with the identical timing with which they were taken, distortion will result, as explained in Figure 10.25. Jitter can be loosely defined as timing errors in analog-to-digital and digital-to-analog conversion. When the analog voltage is reconstructed using a DAC with timing variations in the word clock, the sample amplitudes, the ONEs and ZEROs are correct, but they come out at the wrong time. Because the right amplitude at the wrong time is the wrong amplitude, a timing jitter in the word clock produces an amplitude variation in the DAC output, causing the waveform to change shape. This in turn introduces either spurious components related to the jitter frequency or raises the noise floor of a DAC, unless the jitter is periodic. If the jitter has a Gaussian distribution with a root mean square jitter of Δt, the worst-case SNR resulting from this random word clock jitter is

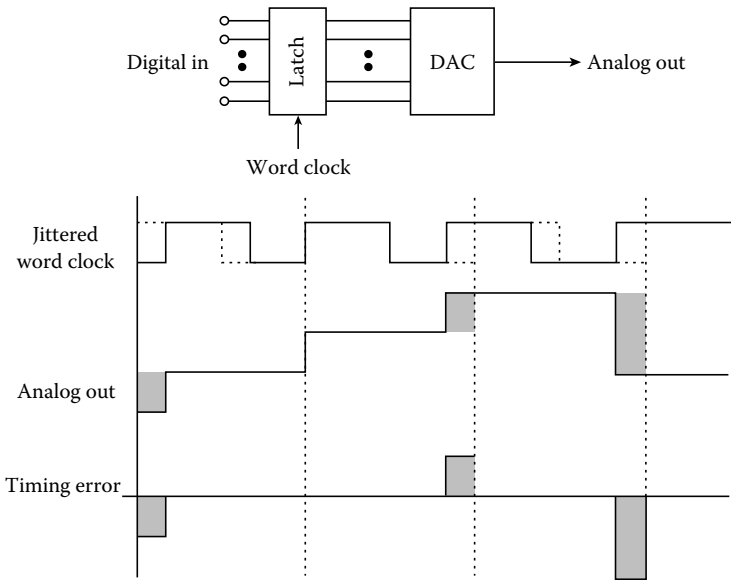


FIGURE 10.25 Word clock jitter effect.

$$\text{SNR} = -20 \times \log \frac{2\pi f \Delta t}{M^{1/2}} \tag{10.18}$$

where f is the signal frequency and M is the OSR. The OSR M is defined as

$$M = \frac{f_c}{2f_n} \tag{10.19}$$

where
f_c is the word clock frequency
f_n is the noise bandwidth

The timing jitter error is more critical in reproducing high-frequency components. In other words, to make an N-bit DAC, an upper limit for the tolerable word clock jitter is

$$\Delta t < \frac{1}{2\pi B 2^N} \left(\frac{2M}{3}\right)^{1/2} \tag{10.20}$$

where B is the bandwidth of the baseband. This implies that the error power induced in the baseband by clock jitter should be no larger than the quantization noise resulting from an ideal N-bit DAC. For example, a Nyquist-sampling 16 bit DAC with a 22 kHz bandwidth should have a word clock jitter of <90 ps.

These timing errors are caused by variation in the clock signal that controls the time when the DAC converts each digital word to an analog voltage. Usually, the term ''clock jitter'' results from everything inside a digital processor (noise from digital circuitry, inductance and capacitance of a clock bus or of a printed circuit board trace) as well as the instability of the clock source. For example, noise in digital circuitry causes the zero-crossing points to shift slightly. If the digital signal has an average slope of 10 V/μs, just 1 mV_{rms} of noise will cause 100 ps of root mean square jitter.

10.1.5.3 Voltage Reference

Ideally, the voltage reference V_{ref} is a constant temperature- and supply-independent voltage with a zero output resistance. The most common voltage reference source is a silicon bandgap voltage reference of about 1.2 V. Depending on the process used, the bandgap reference voltage has a temperature coefficient of typically 20–100 ppm/°C at room temperature when it is set to a voltage ranging from 1.2 to 1.3 V. To generate a different reference voltage other than 1.2 V, op-amps are often used to make inverting and noninverting amplifiers with trimmable feedback gains. Because the bandgap voltage for zero temperature coefficient is not clearly defined and process dependent, it is common to trim this voltage at a wafer level, and it is extremely difficult to achieve an absolute accuracy over a wide range of temperature. However, most DAC applications, except for precision instruments, do not require an absolute accuracy of the reference voltage, and the load-driving capability of the op-amp used in the voltage reference is of paramount interest.

When a DAC is used as a subblock of an ADC, the DAC has only to settle to a final value within a given clock period. In such applications it is not important how the DAC settles. However, the DAC as a stand-along device should have a single-pole response without a glitch and a slew limit. Except for the current-ratioed DAC, the voltage reference is periodically loaded and disturbed by a switched-in load at the clock rate, and the reference output should be restored immediately so that the DAC can settle fast. When disturbed, the voltage reference needs to settle with a time-constant shorter than that of the DAC with an unlimited slew rate. Figure 10.26 illustrates this situation when the voltage reference periodically refreshes the loading capacitor. If the op-amp unity-gain bandwidth is $1/\tau_0$ rad/s, the op-amp restores the output with a time constant $(1 + R_2/R_1)\,\tau_0$ of the feedback amplifier. The large capacitor C_D helps to prevent the op-amp from slewing because the voltage dip of $V_{\text{ref}} C_L/(C_D + C_L)$ decreases as C_D increases.

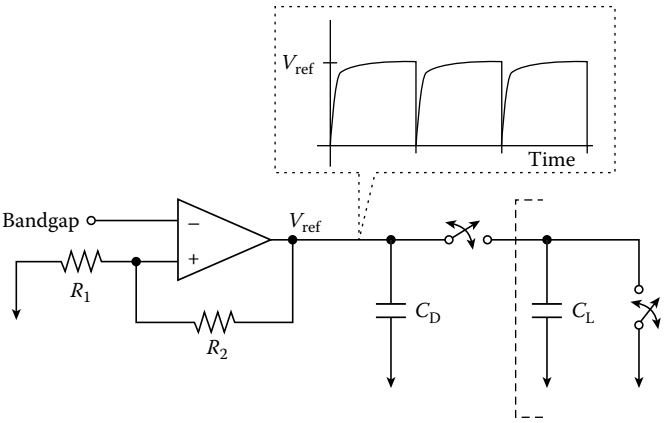


FIGURE 10.26 Voltage reference periodically charging a load capacitor.

10.1.5.4 Noise

Noise is a fundamental limit in high-resolution DACs. The resolution of a DAC is limited by a quantization noise given by Equation 10.2, unless the circuit contributes higher noise. There are many noise sources in real DACs. The DAC output is corrupted by noises directly coupled from bouncing power supplies, bias lines, and ground, as well as the noise sources such as thermal noise, flicker noise, and shot noise. To reduce the former noise sources, it is necessary to carefully separate analog and digital supplies and to eliminate sets of ground loops using a star-ground configuration. The use of a fully-differential architecture helps to reduce such coupling noises. However, the latter set of noise sources is predictable and can be reduced by a low-noise design.

The dominant noise source of a resistor-string DAC shown in Figure 10.4 is a white thermal noise of resistors. If the output resistance of the DAC is R_{out}, the root mean square output noise of the DAC is $(4kTR_{out})^{1/2}$ V/Hz$^{1/2}$, where k is the Boltzman constant and T is the absolute temperature. The thermal noise is about 4 nV/Hz$^{1/2}$ if R_{out} is 1 kΩ. The output resistance of this DAC depends on the digital input code. The worst-case R_{out} is one quarter of the total resistance of the string when the output is connected to the center tap. For lower noise, the total resistance should be minimized at the cost of high power consumption. The noise of the voltage reference appears at the output after being divided by the resistor string. The voltage noise of the output buffer, if used, is directly added to this noise.

On the other hand, the noise source of the current-ratioed DAC of Figures 10.6 and 10.7 is the current noise contributed either by the shot noise of the current source or by the parallel resistor connected to the output node. If the total output current is I_{out} and the total output shunt resistance is R_p, the root mean square output shot noise is 2 qI_{out} A/Hz$^{1/2}$ and the root mean square current noise contributed by the shunt resistor is 4 kT/R_p A/Hz$^{1/2}$, where q is 1.6×10^{-19} C. The shot noise of a 50 μA current and the current noise of a 1 kΩ shunt resistor give the same root mean square noise of 4 pA/Hz$^{1/2}$. The shunt feedback resistance of the transresistance amplifier also contributes to the total noise. If the current DAC is terminated by a low-impedance source of 50–75 Ω, the current noise is usually dominated by this termination resistor.

The main source of the capacitor-array DAC shown in Figures 10.8 and 10.9 is a well-known kT/C noise. The sampled root mean square noise voltage of the capacitor of C is $(kT/C)^{1/2}$ V without regard to the switch-on resistance. The 1 pF sampling capacitor gives a 64 μV sampled noise. This noise is evenly distributed over the Nyquist band (half the sampling frequency), and puts a lower limit on the smallest signal the DAC can handle. Furthermore, the reference and ground noises are divided by the C divider and appear at the output of the DAC. When applied to the oversampling converters either in delta-sigma modulators or in a 1 bit DAC, shown in Figures 10.22 and 10.23, the kT/C noise in the signal band is lower than the Nyquist-rate converters by the OSR of M given by Equation 10.19.

10.1.6 Low-Spurious DAC Design Examples

Modern broadband digital communication systems such as cable modem, cellular baseband, and wireless local area network consist of multiple channels. For high bit rate transmission of 54 Mbit/s over 20 MHz bandwidth, the RF standards such as IEEE 802.11a and Hyperlan2 use the orthogonal frequency division multiplexing scheme that divides the signal band into 52 of 300 kHz narrowband channels. It is challenging to preserve signal integrity of such a complex waveform. Multiple channels are mixed in many different ways when analog signal processing is nonlinear. For example, the multichannel spectrum is perfect when digitally generated, but it is degraded while being converted into analog waveform due to the DAC nonlinearity. To generate an accurate multichannel spectrum without mixing, DAC should be linear both in static and dynamic performance. High spurious-free dynamic range is the most important for frequency-domain applications. Recently, a few techniques such as spatial averaging [13–15], self-trimming [16], and dynamic linearity enhancement [16,17] are reported.

10.1.6.1 Spatial Averaging for Static Linearity

DAC static linearity measured in terms of DNL and INL depends on how DAC components are arranged. In most stand-alone DACs, an array of current sources is predominantly used. It is well known that the binary-weighted array is simple to decode but exhibits poor DNL, while the thermometer-coded array is complex to decode but exhibits good DNL. In fact, the complexity of the thermometer-coded current DAC grows exponentially. Therefore, in typical applications above 8–10 bits, the MSB part is coded in thermometer while the LSB part is coded in binary. The thermometer-coded DAC array occupies a large area, and current source matching is limited by overall gradient effect. In a large thermometer-coded array, systematic and graded errors are far greater in effect than random device mismatch errors. They result from many factors such as edge effect, current source output resistance, supply voltage drop, thermal gradient, and process gradients of doping and thickness across the die, etc. In integrated circuit design, the common-centroid layout scheme has been used to effectively match devices when limited by the gradient effect.

Since the thermometer-coded array is physically large, a two-dimensional common-centroid scheme should be used. To improve INL, the two-dimensional array can be split into four quadrants biased separately using common-centroid bias circuits [13]. To improve INL further, each current source can be split into four small slices connected in parallel, and each slice is placed in one of the four different quadrants [14]. This results in the first-order spatial averaging of the systematic and gradient errors. In general, the gradient effect is first-order, but the second-order effect limits DAC INL performance at 14 bit level or higher. Further partitioning the unit current source into 16 small slices connected in parallel can remove the residual gradient effect [15]. Unlike the first-order centroid scheme, the second-order spatial averaging scheme uses four common-centroid slices placed in four different quadrants. In theory, the matching improves as the array is split into many smaller pieces and the common-centroid scheme is used in each small piece. However, the complexity grows significantly, and the operating speed is inevitably reduced.

10.1.6.2 Self-Trimming for Static Linearity

The two-dimensional geometric averaging is very costly in terms of chip area and the operating speed as the array gets larger. Calibration techniques as discussed are effective in overcoming device mismatch and gradient effects in IC process. One recent trend is self-trimming. The concept is to adjust the current source values electronically like laser trimming. Although laser trimming is a one-time adjustment, the self-trimming is adaptive and continuously engaged. A trimmable floating current source is shown in Figure 10.27, which can be used as one MSB current source in the thermometer-coded array [16]. The floating current source made of transistors N_4 and P_2 is switched from the top side by transistors N_6 and N_7. When the current source needs trimming, the bottom transistor N_2 is turned on, and the current is switched to the measuring resistor R_2. This arrangement of trimming is transparent to the normal

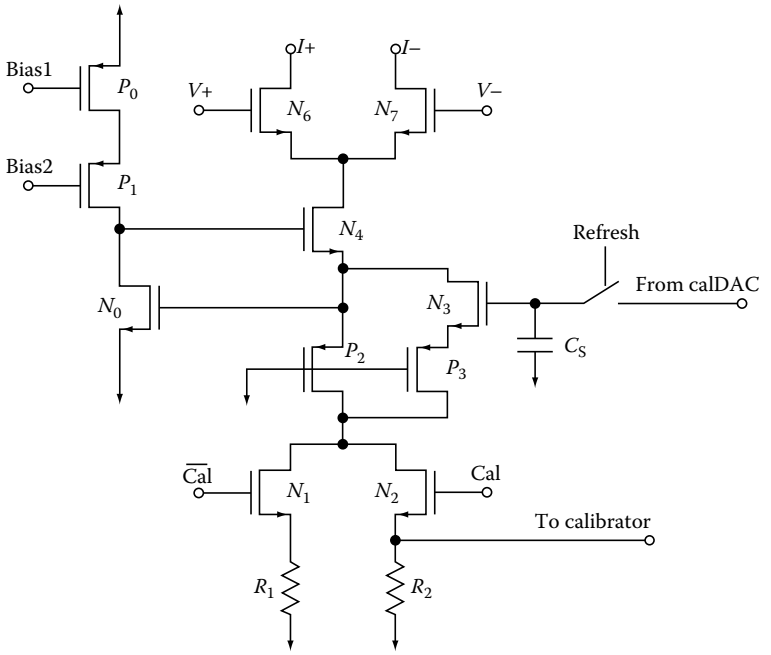


FIGURE 10.27 Trimmable floating current source.

operation and is performed in background. An oversampling calibrator based on delta-sigma modulation repeatedly monitors the voltage across R_2, accurately compares it with the preset reference voltage, and constantly feedbacks the error signal that periodically updates the capacitor C_s to adjust the current source.

10.1.6.3 Dynamic Linearity Enhancement

The static linearity up to over 14 bits can be achieved either with spatial averaging or with self-trimming, but the dynamic linearity heavily depends on how the DAC output settles. For example, the wireless baseband spectrum covers over 20 MHz bandwidth, but generating such a wideband signal with high static and dynamic linearity is not a trivial task. Furthermore, in scaled CMOS, devices are more nonlinear and low supply voltage often limits the DAC linearity. In standard current DACs, each current source is switched to the output at different times due to clock skew, and adds parasitic capacitance to the output node, which results in code-dependent output settling. The clock skew and the code-dependent settling are most prominent among many factors limiting dynamic DAC linearity. Isolating the output node from current sources and synchronizing the code transitions are simple solutions for these, but the clock skew is difficult to deal with high precision. It is often suggested that the DAC's dynamic linearity can be improved if its output is sampled and held using a deglitcher. In fact, the deglitcher is functionally a track and hold circuit, but it is also difficult to implement at high frequencies with high precision.

The return-to-zero (RZ) scheme solves the deglitcher problem as explained in Figure 10.28a [17]. While the error in the track and hold stays for the half period, the error in the RZ exponentially decays. The logic behind this is that rather than sampling analog voltage accurately, it is easier to let the output keep on changing exponentially. Exponential settling is a linear process. The same effect as RZ can be achieved in a simpler manner, called track and attenuate [16]. Rather than resetting the DAC output, just attenuating it achieves the same goal. Figure 10.28b illustrates the DAC output stage performing the track and attenuate function. The differential DAC output current comes into the folded-cascode node V_x, which is isolated from the output node. The output nodes track the signal during the half clock period,

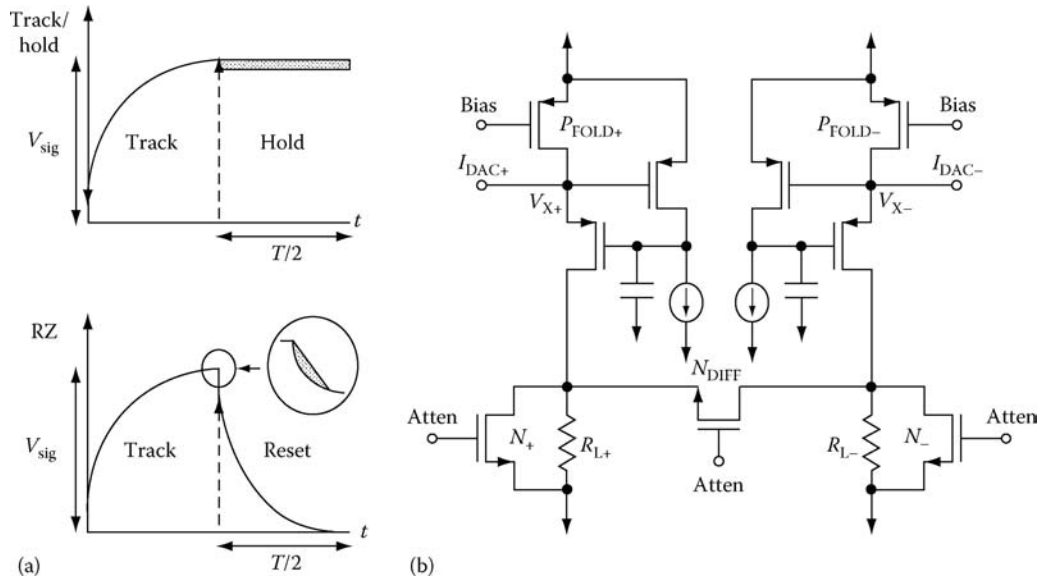


FIGURE 10.28 (a) Track and hold vs. RZ and (b) track/attenuate output stage.

but are bypassed by three transistors N_+, N_-, and N_{DIFF} for attenuation during the remaining half clock period. The difference of the RZ DAC from the conventional DAC is that the output spectrum nulls at multiples of twice the sampling frequency, which in turn provides less attenuation in the passband.

References

1. D. M. Freeman, Slewing distortion in digital-to-analog conversion. *J. Audio Eng. Soc.*, 25, 178–183, 1977.
2. R. B. Craven, An integrated circuit 12-bit D/A converter. *Dig. IEEE Int. Solid-State Circuits Conf.*, San Francisco, CA, pp. 40, 41, 1975.
3. D. T. Comer, A monolithic 12-bit DAC. *IEEE Trans. Circuits Syst.*, CAS-25, 504–509, 1978.
4. J. M. McCreary and P. R. Gray, All-MOS charge redistribution analog-to-digital conversion techniques I. *IEEE J. Solid-State Circuits*, SC-10, 371–379, 1975.
5. R. J. Van de Plassche, Dynamic element matching for high accuracy monolithic D/A converters. *IEEE J. Solid-State Circuits*, SC–11, 795–800, 1976.
6. M. J. M. Pelgrom and M. Roorda, An algorithmic 15-bit CMOS digital-to-analog converter. *IEEE J. Solid-State Circuits*, SC-23, 1402–1405, 1988.
7. D. W. J. Groeneveld, H. J. Schouwenaars, H. A. H. Termeer, and C. A. A. Bastiaansen, A self-calibration technique for monolithic high-resolution D/A converters. *IEEE J. Solid-State Circuits*, SC-24, 1517–1522, 1989.
8. J. C. Candy and A. N. Huynh, Double interpolation for digital-to-analog conversion. *IEEE Trans. Commun.*, 34, 77–81, 1986.
9. H. S. Lee, D. A. Hodges, and P. R. Gray, A self-calibrating 15 bit CMOS A/D converter. *IEEE J. Solid-State Circuits*, SC-19, 813–819, 1984.
10. S. H. Lee and B. S. Song, Digital-domain calibration of multistep analog-to-digital converters. *IEEE J. Solid-State Circuits*, SC-27, 1679–1688, 1992.
11. T. Hayashi, Y. Inabe, K. Uchimura, and T. Kimura, A multistage delta-sigma modulator without double integration loop. *Dig. IEEE Int. Solid-State Circuits Conf.*, pp. 182, 183, 1986.

12. W. L. Lee and C. G. Sodini, A topology for higher-order interpolative coder. *Proc. Int. Symp. Circuits Syst*., pp. 459–462, 1987.
13. C. H. Lin and K. Bult, A 10-b, 500-MSample/s CMOS DAC in 0.6 mm^2. *IEEE J. Solid-State Circuits*, SC-33, 1948–1958, 1998.
14. J. Bastos, A. M. Marques, M. S. J. Steyaert, and W. Sansen, A 12-bit intrinsic accuracy high-speed CMOS DAC. *IEEE J. Solid-State Circuits*, SC-33, 1959–1969, 1998.
15. G. A. M. Van der Plas, J. Vandenbussche, W. Sansen, M. S. J. Steyaert, and G. G. E. Gielen, A 14 bit intrinsic accuracy Q^2 random walk CMOS DAC. *IEEE J. Solid-State Circuits*, SC-34, 1708–1718, 1999.
16. A. R. Bugeja and B. S. Song, A self-trimming14-b100-MS/s CMOS DAC. *IEEE J. Solid-State Circuits*, SC-35, 1841–1852, 2000.
17. A. R. Bugeja, B. S. Song, P. L. Lakers, and S. F. Gillig, A 14-b, 100MS/s CMOS DAC designed for spectral performance. *IEEE J. Solid-State Circuits*, SC-34, 1719–1732, 1999.

10.2 Analog-to-Digital Converters

Ramesh Harjani

10.2.1 Introduction

With the increased complexity possible in modern-day integrated circuits, analog-to-digital converter (ADC) and digital-to-analog converter (DAC) have become ubiquitous components of mixed-signal integrated circuits. ADCs transform an analog signal, V_A, into an N-bit digital representation, V_d. Such a converter is said to have a resolution of N bits. The digital signal V_d is an approximation of the original analog signal, V_A, and the maximum error during this conversion process for an N-bit converter is equal to $1/2^N$ of the full-scale value. This error is called the quantization error.

A number of topologies exist for ADC. They can be classified as Nyquist rate converters or oversampled converters. Nyquist rate converters sample the input at the minimum sampling rate, i.e., two times the maximum signal frequency. As the name implies, oversampled converters take more samples than is mandated by the Nyquist criterion to generate extremely high resolution. Nyquist rate converters are usually classified as (1) high-speed, (2) medium-speed, and (3) high-resolution converters. However, the different architectures are better categorized by evaluating the number of clock cycles they use to perform the analog-to-digital conversion. For example, for N bits of resolution a high-speed converter performs the conversion in one or two clock cycles, a medium-speed converter performs the conversion in N clock cycles, and a high-resolution converter performs the conversion in 2^N clock cycles. Thus, we classify them here as 1-, N-, and 2^N-clock converters.

Before describing the details of the various ADC architectures we discuss some of the performance characteristics of ADCs in general and describe some test techniques that are used to measure these characteristics.

10.2.1.1 ADC Test Techniques

ADCs are primarily tested using parametric techniques. That is, a key set of parameters that characterize an ADC are verified. An ADC is characterized by its static and dynamic performance. In static performance, we are primarily concerned about the linearity of the I–O transfer characteristics, and in dynamic performance we are concerned about the operation of the converter at full operating speed.

Figure 10.29 shows the transfer characteristics for an ideal 3 bit converter. The dashed line shows the transfer characteristics of an infinite precision converter and the bold line shows the transfer characteristics of a 3 bit version. We note that the least significant bit (LSB) value is equal to $1/2^N$ of the full-scale value. Figures 10.30 through 10.33 show examples of the static performance characteristics of an ADC.

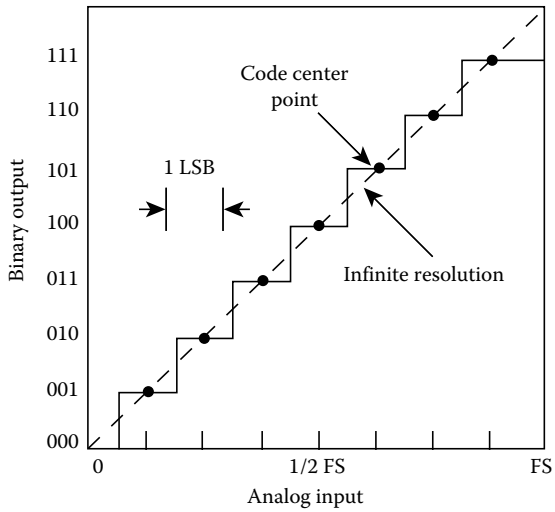


FIGURE 10.29 Transfer characteristics for an ideal 3 bit ADC.

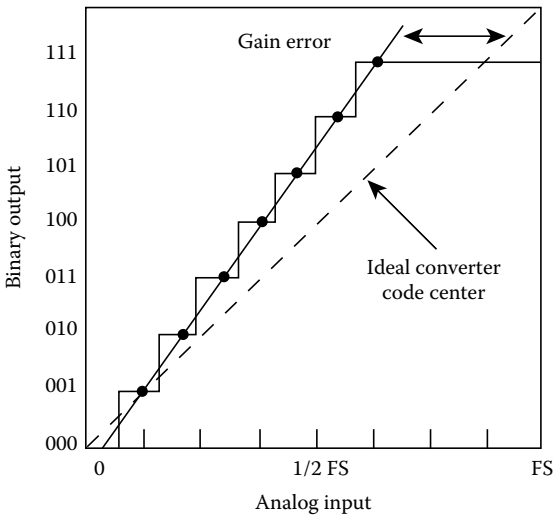


FIGURE 10.30 Gain error in an ADC.

A gain error is said to be present when the maximum digital value does not correspond to the full-scale analog value. An offset error corresponds to a horizontal shift in the transfer characteristics. The integral nonlinearity error specifies the maximum deviation of the transfer characteristics from the ideal code center. Differential nonlinearity specifies the deviation of each stepsize from $1/2^N$ of the full-scale value.

Dynamic ADC performance characteristics include signal-to-noise ratio (SNR), effective bits, aperture errors, and input signal bandwidth. During full-speed operation some additional errors become evident because of the finite settling time and bandwidth limitations of the circuits within the ADC. The SNR and effective bits are dynamic ways of measuring the minimum resolution and errors in the transfer characteristics. Different terms are used to specify similar performance parameters largely because different measurement techniques are used to generate them. The input signal bandwidth specifies the maximum frequency of the input signal.

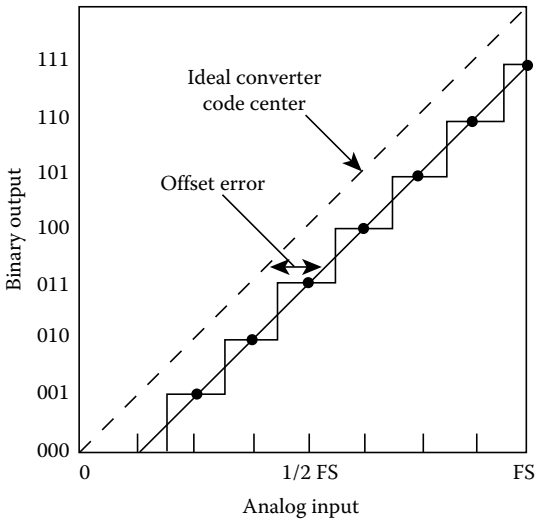


FIGURE 10.31 Offset error in an ADC.

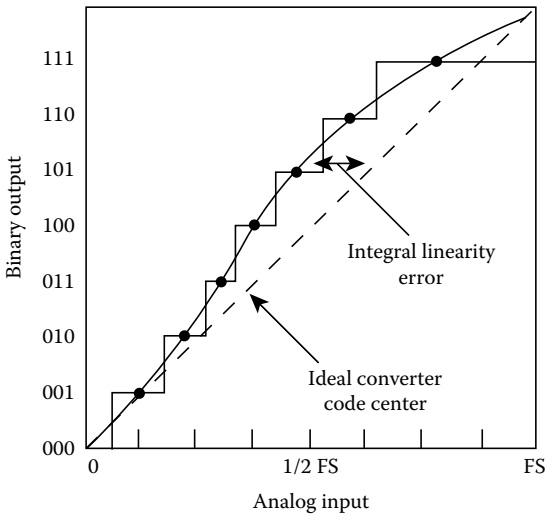


FIGURE 10.32 Integral nonlinearity error.

Because of the large number of performance specifications for ADCs, quite a few techniques are used to test them as well. These testing techniques also reflect the different types of performance characteristics, i.e., static and dynamic. The more traditional methods of testing ADC are primarily concerned with checking the static characteristics of these converters. Examples of such test techniques include analog difference signal methods, crossplot methods, and servo loop code transition measurement methods [3]. Dynamic techniques to test linearity are based on using a well-known and near-full-scale input signal and evaluating the output code probability over a large input sample size. This technique is called the ''code density test'' or the ''histogram test.'' The output code probability for a linear ramp input is uniform for all codes. However, it is difficult to generate extremely accurate ramp signals at high speed, therefore, sine-wave signals are used. Unfortunately, with a sine wave the output code probability is no longer uniform but is instead ''cusp-shaped''. An estimate of the differential nonlinearity can be

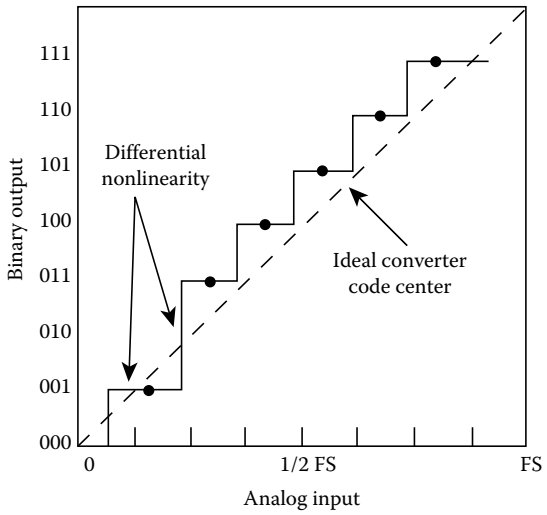


FIGURE 10.33 Differential nonlinearity error.

generated by evaluating the difference between the expected code probability and measured values. The code density test requires a large number of data samples in the range of several hundreds of thousands. Additionally, a small number of large magnitude errors are easily masked in this technique. More recently, frequency domain-based techniques have been used to measure the harmonic content of the converted signal to provide an estimate of the SNR and effective number of bits. In this technique, a discrete time Fourier transform (DTFT) is performed on the output data sequence and is used to measure the data converter performance characteristics. Although the number of data samples required are fewer than the code density approach, this approach still requires a few thousand data samples. Additionally, proper windowing functions and synchronized sampling may be required to reduce spectral leakage.

Data converters are relatively time-consuming to test. However, rather than explain all the techniques that can be used to test data converters, we concentrate on a single method, the crossplot method, to provide the reader with some insight into the complexity of testing ADCs. For additional details about the other methods, readers are referred to [3]. A block diagram for the crossplot method is shown in Figure 10.34. In the crossplot technique, the output of the lowest 2 or 3 bits of an ADC are fed to the

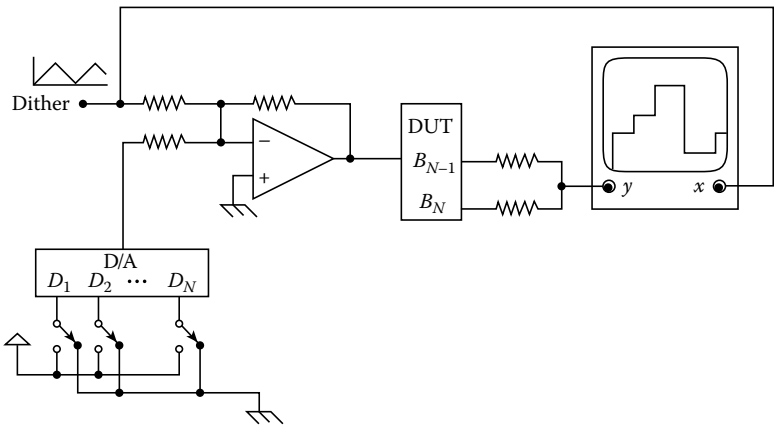


FIGURE 10.34 Crossplot technique used to test ADC static characteristics. (From Demler, M.J., *High-Speed Analog-to-Digital Conversion*, Academic Press, New York, 1991, 162.)

Y input of an oscilloscope and a separate triangular dither signal is fed to the *X* input of the oscilloscope. The input to the ADC is generated using a discrete summing amplifier that adds the output of a DAC, which is itself swept through its input space at a much slower rate, and the dither signal, mentioned earlier. This technique generates a staircase waveform on the oscilloscope. Therefore, the linearity of the last 2 or 3 bits around a fixed bias voltage, which is generated by the DAC, can be seen easily on an oscilloscope. The primary advantage of this technique is that it uses fairly inexpensive equipment. However, it is extremely time consuming to evaluate all 2^N combinations and the technique provides no information about the integral linearity, offset, and gain errors. In general, techniques that are used to test ADCs are fairly complex and either require a large sample size, i.e., high cost, or use extremely complex equipment to perform the tests.

Having described the basic characteristics of ADC, we now describe of the different ADC architectures in detail. We first consider Nyquist rate converter topologies and then describe oversampled converters.

10.2.2 Nyquist Rate Converters

As mentioned earlier, Nyquist rate converters are classified as 1-clock converters, *N*-clock converters, and 2^N-clock converters. Examples of 1-clock converters include the flash architecture, the pipelined architecture [11,12], and the voltage folding architecture. Examples of *N*-clock converters include the successive-approximation architecture [9] and the algorithmic architecture [6]. Examples of 2^N-clock converters include the single-slope and dual-slope architectures [4].

10.2.2.1 1-Clock Converters

Flash converters. Flash or parallel converters are the highest rate converters and are sometimes also called ''video rate converters'' because they operate at rates necessary for video signals. Parallel converters are O(1)-clock converters in that they require one or two clock cycles to perform an *N*-bit conversion. The basic principle of operation is fairly simple and is easily explained with the help of Figure 10.35. This figure is a block diagram for a 3 bit flash converter. The input is compared to $(2^3 - 1)$ comparators. The comparison voltage for the *i*th comparators is set to be equal to $(iV_{FS})/(2^3)$, where V_{FS} is the full-scale voltage. The resulting outputs are then converted from thermometer code to binary code. In general, the time required to perform the analog-to-digital conversion is equal to the comparator resolution time plus the time taken to perform the digital code conversion, i.e., one clock cycle. The reference voltages for the comparators are usually generated using an equally spaced resistor string that is offset by $R/2$, as shown in Figure 10.36. The primary advantage of providing the $R/2$ offset is that it allows the code center to pass through the origin. The resolution of integrated MOS flash converters is limited to approximately 8 bits. This is primarily due to the matching constraints placed on the resistors and also because of the exponential increase in the number of components ($\approx 2^N$) required for higher resolution. The speed for MOS flash converters is usually limited to about 50 MHz. However, if higher speed operation is required, then the process of time interleaving may be employed. Here, *M* (*N*-bit) ADCs are operated in parallel, but each is delayed somewhat, as shown in Figure 10.37 [1]. By starting the conversion operation for A/D_2 before the operation of A/D_1 is complete allows the complete interleaved converter to operate at *M* times the speed of each individual converter. Although the technique allows for higher operating speed, the number of parallel paths has increased. Therefore, due to size limitations the resolution is usually lower than for simple flash converters.

Subranging converters. The exponential area penalty ($\alpha\ 2^N$) for flash converters can be mitigated by using subranging techniques without incurring severe speed penalties. Here, instead of deciding all *N*-bits at one time, only a subset of these is decided during the first clock phase, and the rest are decided in the following clock phases. If two clock phases are used to decide all *N*-bits, this is called a two-step subranging converter topology.

A block diagram for a two-step subranging converter is shown in Figure 10.38. In this converter topology the α most significant bits (MSBs) are decided by the coarse ADC, the results of which are then

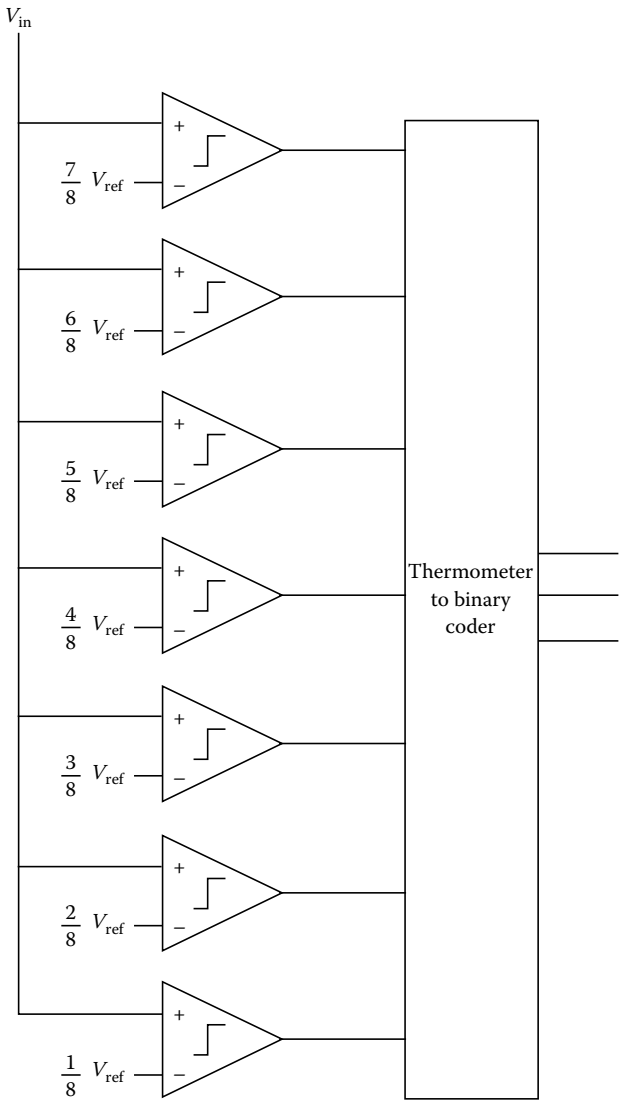


FIGURE 10.35 Three bit flash converter block diagram.

passed on to an α-bit DAC. The output voltage of the DAC is subtracted from the input signal. The resulting voltage is amplified 2α times and then a $(N-\alpha)$-bit fine converter is used to resolve the remaining bits. The entire process—coarse conversion, digital-to-analog subtraction, amplification, and fine conversion—is performed in one clock period. The coarse and fine converters are usually implemented as simple flash converters. Therefore, the operating speed of a pure subranging converter is at least twice as slow as that of a flash converter. However, the total number of comparators (and associated circuits) is reduced from 2^N to $2\alpha + 2^{N-\alpha}$. The savings in area can be substantial. Usually, the value of α is selected to be roughly equal to $N/2$. For example, for $N = 10$ and selecting $\alpha = 5$, the number of comparators reduce from $2^{10} = 1024$ to $2^6 = 64$, which is a savings of 93.75%. This subranging process can be extended further to a larger number of sequential stages; however, the total propagation delay usually limits the operation speed fairly significantly.

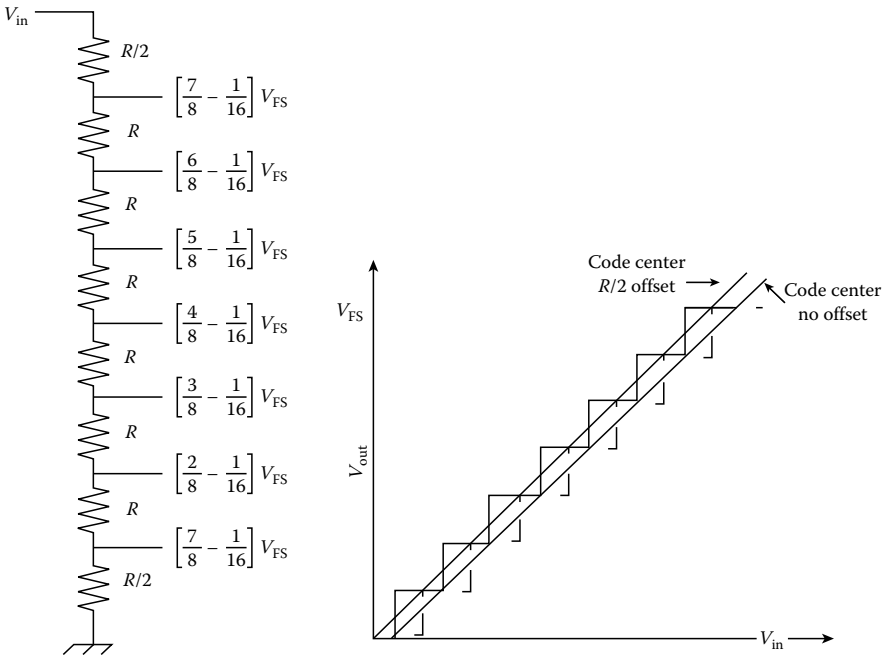


FIGURE 10.36 Flash converter 1/2 LSB offset.

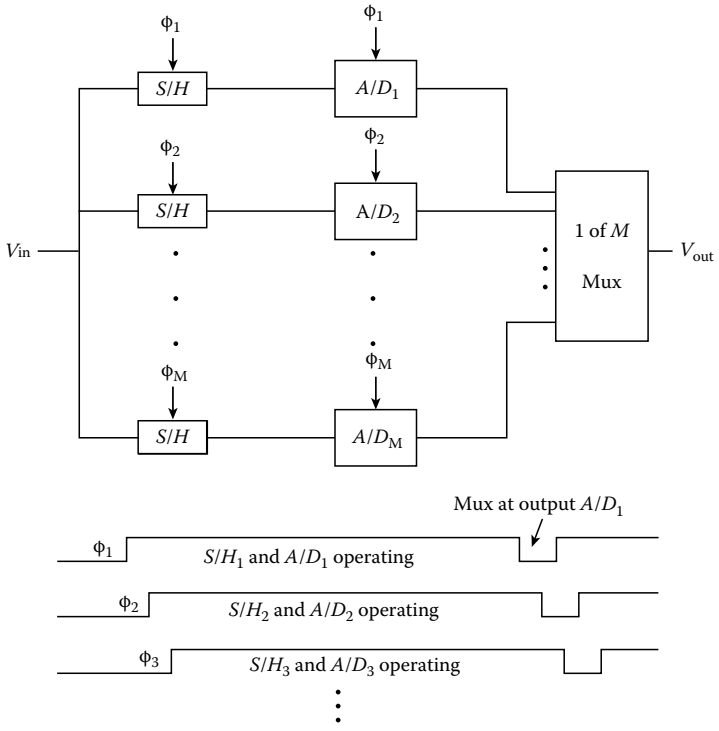


FIGURE 10.37 Time interleaved converter block diagram.

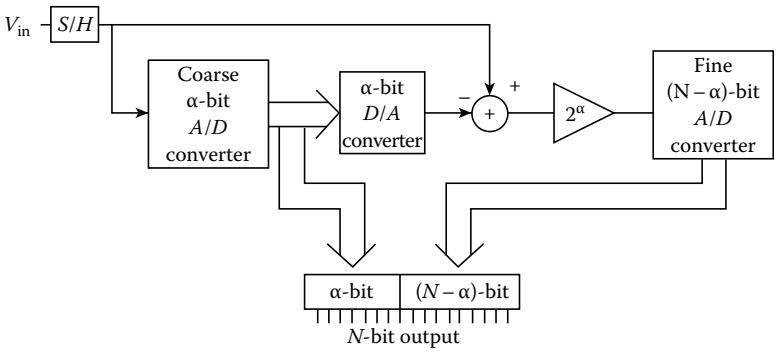


FIGURE 10.38 Subranging converter block diagram.

Pipelined converters. Instead of operating both the coarse and fine converter in the same clock period, a sample-and-hold (S/H) could be added to the circuit in Figure 10.38, such that when the fine converter is resolving the lower $(N-\alpha)$ bits for the first input sample the coarse converter can begin resolving the upper α bits for the next input sample. This kind of converter can be made to operate at speeds comparable to flash converters and is called a pipelined converter. A block diagram for a two-stage pipelined converter is shown in Figure 10.39. Because the α MSBs are decided during the previous clock period, the results from the fine and coarse converters need to be synchronized by delaying the output of the coarse converter by a single clock period. As with the subranging converter the number of comparators have decreased from 2^N to $2^{N-\alpha}+2\alpha$.

Unlike the subranging topology, the pipeline methodology can easily be extended to N sequential stages, each resolving only 1 bit, because of the sample-and-held circuits between each of the stages. The block diagram for such an N-step converter is shown in Figure 10.40. In this figure, note that only one comparator is used per stage and the interstage gain block (also performing the S/H function) has a gain of two. Once again for synchronization purposes, shift registers are used. This pipelined converter [11,12] operates as follows. The input is first compared to see if it is greater than V_{ref}. If the input is greater than V_{ref}, then we set the MSB bit. If the MSB is set, then V_{ref} is subtracted from the input voltage. However, if the MSB is not set, nothing is done. The resulting voltage is then doubled. During the next clock cycle, the resulting voltage is once again compared to V_{ref} to give the MSB-1 bit. While the second stage generates the second MSB for the first input, the first stage processes a new input. When the pipeline is filled, N inputs are being processed simultaneously. Several shift registers are used to ensure that all the data bits

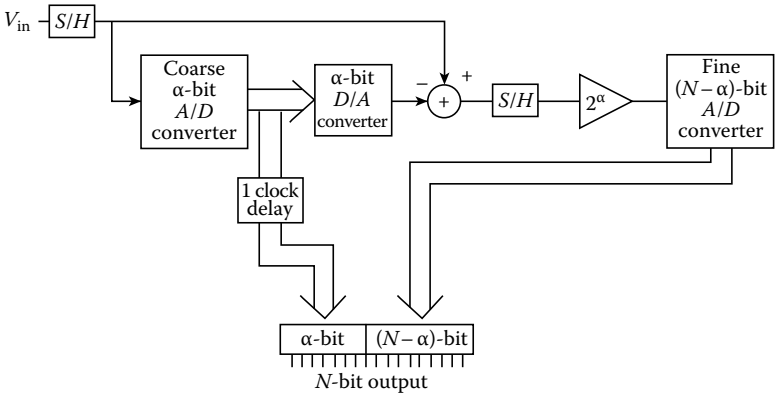


FIGURE 10.39 Two-stage pipelined converter block diagram.

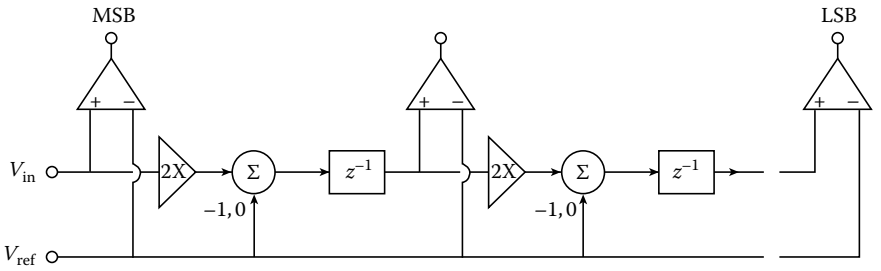


FIGURE 10.40 *N*-stage pipelined converter.

corresponding to a single input are output simultaneously. This process continues until we finally generate the LSB. The analog output of the *i*th stage can be written as

$$V_i = 2[V_{i-1} - b_i V_{\text{ref}}]z^{-1} \tag{10.21}$$

where b_i is equal to 1 if the *i*th bit is set and equal to 0 otherwise.

Figure 10.41 is a simple circuit realization for a pipelined ADC [13]. Other more advanced topologies [12] reduce some of the problems associated with this topology, but operate on the same principle. Only one stage of an *N*-bit converter is shown in this figure. In this first stage, the amplifier A_1 performs the comparison function and the amplifier A_2 performs the 2× multiplication and the conditional subtraction function. Both amplifiers, along with the associated capacitors and switches, also implement a S/H function. Let us first concentrate on the comparison function. For a sufficiently large gain, during ϕ_1 the capacitor C_h is charged to $V_a - V_{off}$, where V_{off} is equal to the amplifier offset voltage. During ϕ_2 the reference voltage, V_{ref}, is connected to the positively charged terminal of capacitor C_h and the feedback around the amplifier is removed. During this period the amplifier A_1 acts as a comparator and compares the voltage $[V_{ref} - (V_a - V_{off})]$ against V_{off}; i.e., the effect of the amplifier offset voltage is completely

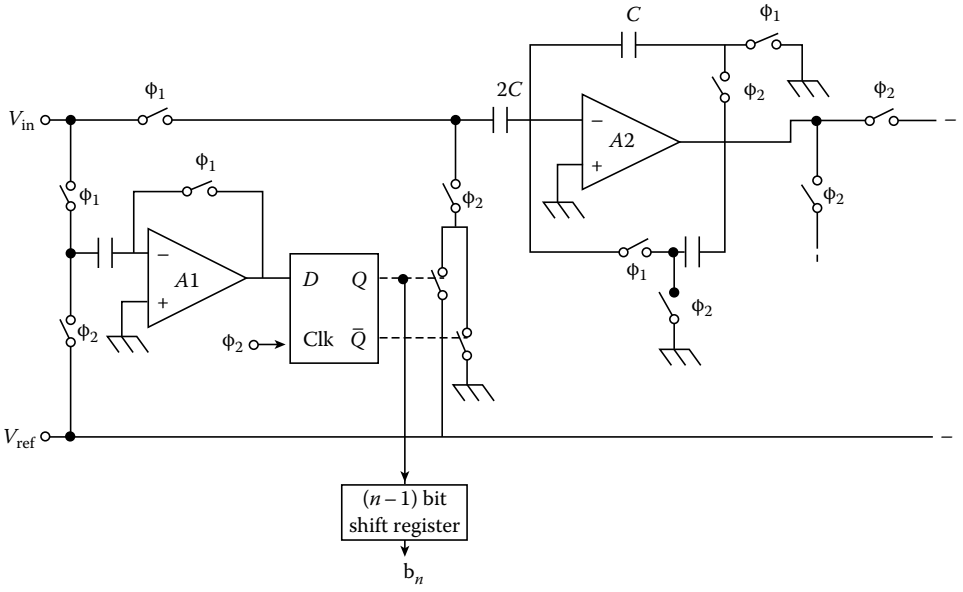


FIGURE 10.41 Circuit realization of an *N*-stage pipelined converter.

canceled. The preceding discussion is valid only for sufficiently large gain. However, if the gain of amplifier A_1 is not sufficiently large then the charge transfer is not complete and an error is introduced in the comparison; i.e., the effective comparison is not with V_{ref} but with $V_{\text{ref}} + \Delta V$. The error in the comparison is dependent on the input signal voltage. Note that the first stage samples the input at ϕ_1 while the next stage samples the input at ϕ_2. Likewise, the third stage samples the input at ϕ_1 and the fourth stage samples the input at ϕ_2, and so forth.

Digital-error correction in multistep converters. Pipelined and subranging converters that have an interstage gain >1 can utilize digital-error correction [14] to improve linearity. We shall illustrate the principle of digital-error correction with the help of a 4 bit two-stage pipelined converter. Each of the stages resolves 2 bits. Digital-error correction can be used to correct for linearity errors in all except the last stage. Additionally, it is unable to correct for digital-to-analog linearity and op-amp settling time errors. Therefore, for our two-stage example we will only be able to correct for errors in the first stage and we shall assume an ideal DAC.

Figure 10.42 is the block diagram for the 4 bit two-stage pipelined converter without digital-error correction. This circuit is a 4 bit version of Figure 10.39. The input signal is sample-and-held by S/H_1. The course MSB bits for the overall converter are generated by the first-stage subconverter (A/D_1). The analog value corresponding to these bits is then generated by the first-stage digital-to-analog subconverter (D/A_1). The difference between the input signal and the digital-to-analog output is called the residue. This residue is amplified by the interstage gain stage ($G = 4$) and passed on to the second-stage analog-to-digital subconverter. The second-stage analog-to-digital subconverter (A/D_1) then generates the lower 2 bits. Because the second stage is working on the signal after one clock delay, an intermediate delay stage is added to synchronize the outputs of the two stages.

The residue for an ideal converter varies from $-\frac{1}{2}$ LSB to $+\frac{1}{2}$ LSB of the first-stage subconverter resolution, as shown in Figure 10.43. In the case of nonlinearity in the first subconverter the residue will have excursions above and below the $\pm\frac{1}{2}$ LSB value, as shown in Figure 10.44. For an ideal digital-to-analog conversion, the residue corresponding to each digital code is still accurate and no data have been lost as yet. In the traditional pipelined converter shown in Figure 10.42 any residue value from the first converter that is greater than $\pm\frac{1}{2}$ LSB of the first stage saturates the second stage and produces errors.

If, however, we change the overall converter topology such that the resolution of the second subconverter is increased by 1 bit, i.e., we double the number of levels, and reduce the interstage gain by half, then we can detect when the residue exceeds the $\pm\frac{1}{2}$ LSB levels and correct for its effect digitally. Figure 10.45 is a block diagram for the pipelined converter in Figure 10.42 with digital-error correction. Whenever the residue from the first stage exceeds $+\frac{1}{2}$ LSB it implies that the digital output of the first

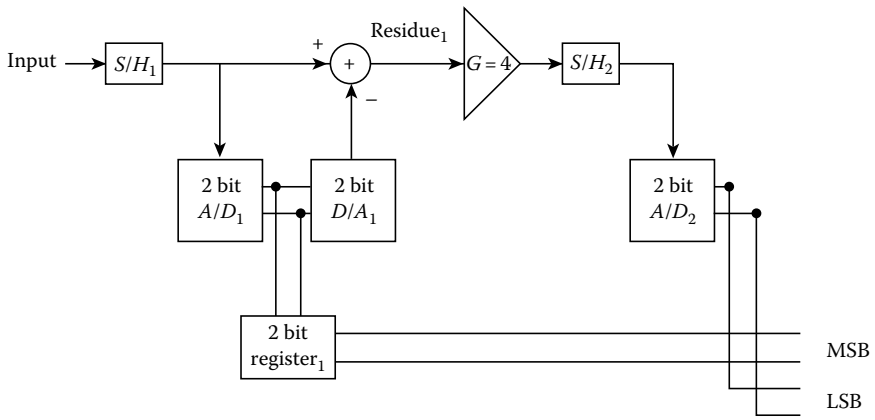


FIGURE 10.42 Four bit, two-stage, pipelined converter block diagram without error correction.

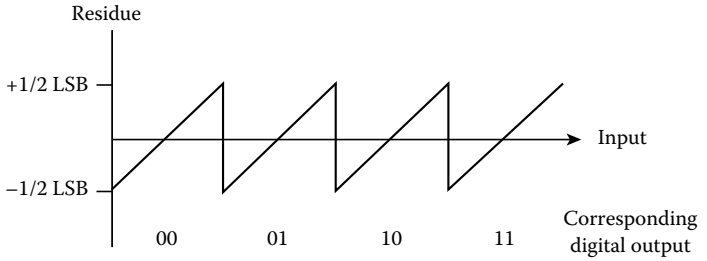


FIGURE 10.43 Ideal subconverter residue.

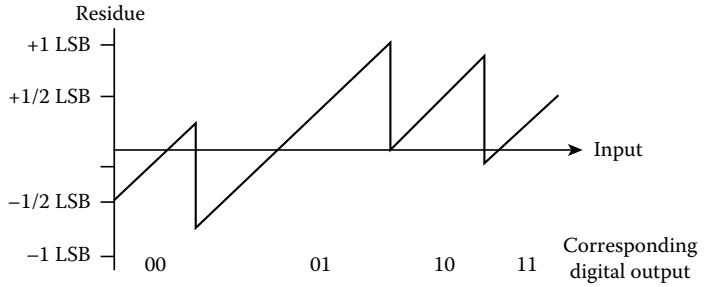


FIGURE 10.44 Nonideal subconverter residue.

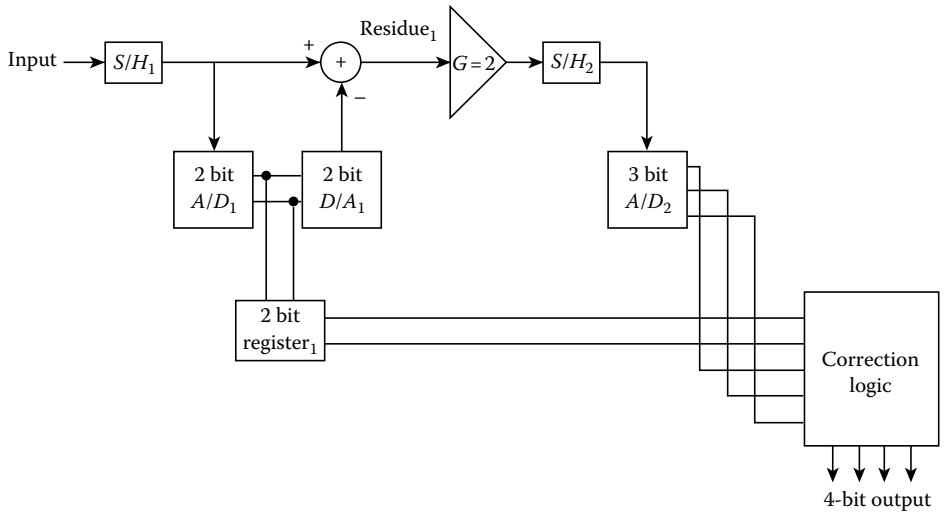


FIGURE 10.45 Four bit, two-stage, pipelined converter block diagram with error correction.

stage subconverter is too small. Likewise, whenever the residue is less than $-\frac{1}{2}$ LSB it implies that the digital output of the first stage is too large. By adding a $\frac{1}{2}$ LSB offset at the input of the first analog-to-digital subconverter and at the output of the first DAC the input to the second subconverter for an ideal first subconverter is restricted between $\frac{1}{4}$ full-scale and $\frac{3}{4}$ full-scale. Any excursion outside this region implies an error in the first analog-to-digital subconverter. The approximate value for this error is measured by the second-stage subconverter and is then subtracted digitally from the final value.

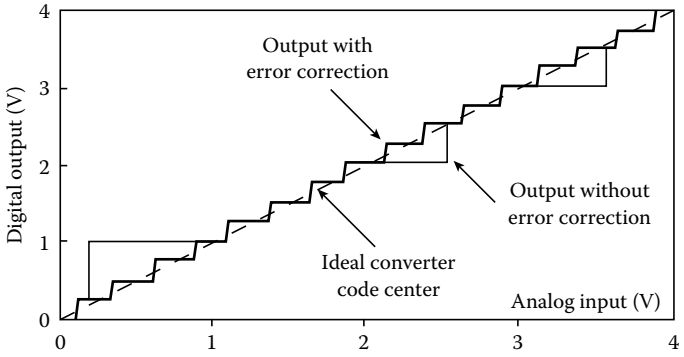


FIGURE 10.46 Digital-error correction simulation results.

Nonlinearities from the second stage are not corrected via this scheme; however, as the interstage gain is greater than 1 the effect of the nonlinearities in the second stage will have a much lower effect than those resulting from the first stage. Nonlinearities in the DAC can be reduced substantially by utilizing reference feedforward compensation [15]. Here, the reference for the second stage changes dynamically and is obtained by amplifying the first-stage digital-to-analog subconverter segment voltage that corresponds to the most current digital output code of the first-stage analog-to-digital subconverter. Figure 10.46 presents the simulation results for the 4 bit two-stage pipelined converter with and without digital error correction. For purposes of clarity the second-stage subconverter is made ideal. In a real converter some nonlinearity would still exist, but would be limited to that introduced by the last stage. Traditionally, even though the resolution of the first subconverter is only 2 bits, it needs to be linear to the overall converter resolution. Digital-error correction can reduce the linearity requirements such that it is commensurate with its resolution.

10.2.2.2 *N*-Clock Converters

Both the successive approximation and algorithmic analog-to-digital topologies require N clock cycles to perform an N-bit conversion. They both perform 1 bit of conversion per clock cycle. The successive approximation converter is a subclass of the subranging converter, in which during each clock cycle only 1 bit of resolution is generated. The algorithmic converter is a variation of the pipelined converter, in which the pipeline is folded back into a loop. Both topologies essentially perform a binary search to generate the digital value. However, in the case of the successive approximation converter the binary search is performed on the reference voltage, while in the case of the algorithmic converter the search is performed on the input signal.

Successive approximation converters. A block diagram for the successive approximation converter is shown in Figure 10.47. Because the conversion requires N clock cycles a S/H version of the input signal is provided to the negative input of the comparator. The comparator controls the digital logic circuit that performs the binary search. This logic circuit is called the successive approximation register (SAR). The output of the SAR is used to drive the DAC that is connected to the positive input of the comparator.

During the first clock period, the input is compared with the MSB, i.e., the MSB is temporarily raised high. If the output of the comparator remains high, then the input lies somewhere between 0 and $V_{ref}/2$, and the MSB is reset to 0. However, if the comparator output is low, then the input signal is somewhere between $V_{ref}/2$ and V_{ref}, and the MSB is set high. During the next clock period the MSB-1 bit is evaluated in the same manner. This procedure is repeated such that at the end of N clock periods all N-bits have been resolved. Figure 10.48 is the binary search procedure for a 4 bit converter and shows the comparator output sequence that corresponds to an input equal to 72% of V_{ref}.

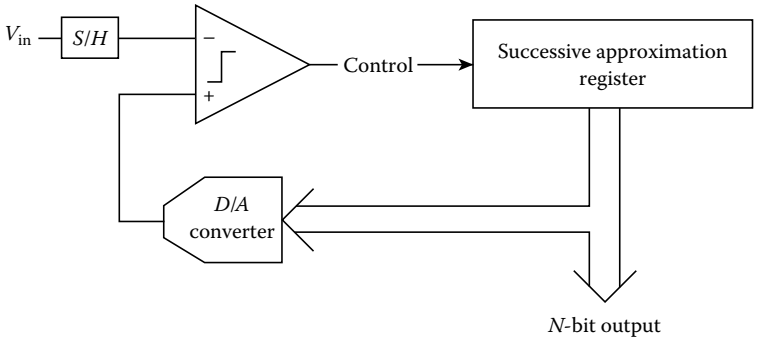


FIGURE 10.47 Successive approximation converter block diagram.

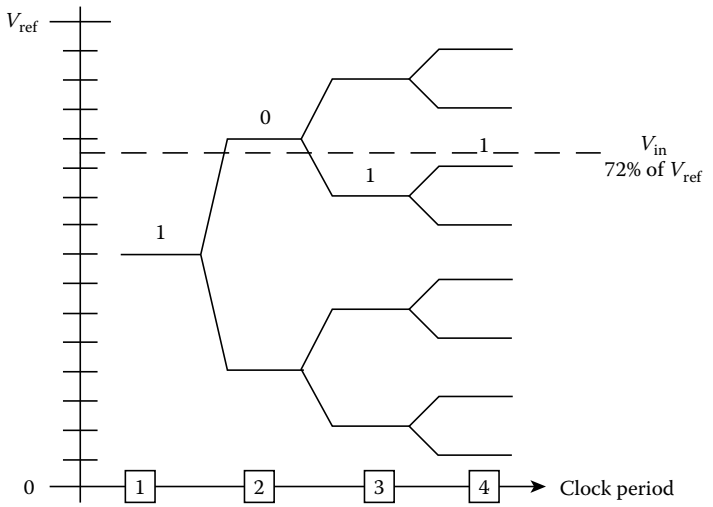


FIGURE 10.48 Binary search process for successive approximation.

The successive approximation converter is one of the most popular topologies in both MOS and bipolar technologies. In MOS technologies the charge-redistribution implementation [9] of the successive approximation methodology is the most commonly used. The circuit diagram of a 4 bit charge redistribution converter is shown in Figure 10.49. In this circuit, the binary weighted capacitors $\{C, C/2, \ldots, C/8\}$ and the switches $\{S_1, S_2, \ldots, S_5\}$ form the 4 bit scaling DAC. For each conversion the circuit operates as a sequence of three phases. During the first phase (sample) switch S_0 is closed and all the other switches $\{S_1, S_2, \ldots, S_6\}$ are connected such that the input voltage V_{in} is sampled onto all the capacitors. During the next phase (hold) S_0 is open and the bottom plates of all the capacitors are connected to ground; i.e., switches $\{S_1, S_2, \ldots, S_5\}$ are switched to ground. The voltage, V_x, at the top plate of the capacitors at this time is equal to $-V_{in}$ and the total charge in all the capacitors is equal to $-2CV_{in}$. The final phase (redistribution) begins by testing the input voltage against the MSB. This is accomplished by keeping the switches $\{S_1, S_2, \ldots, S_5\}$ connected to ground and switching S_1 and S_6 such that the bottom plate of the largest capacitor is connected to V_{ref}. The voltage at the top plate of the capacitor is equal to

$$V_x = \frac{V_{ref}}{2} - V_{in} \tag{10.22}$$

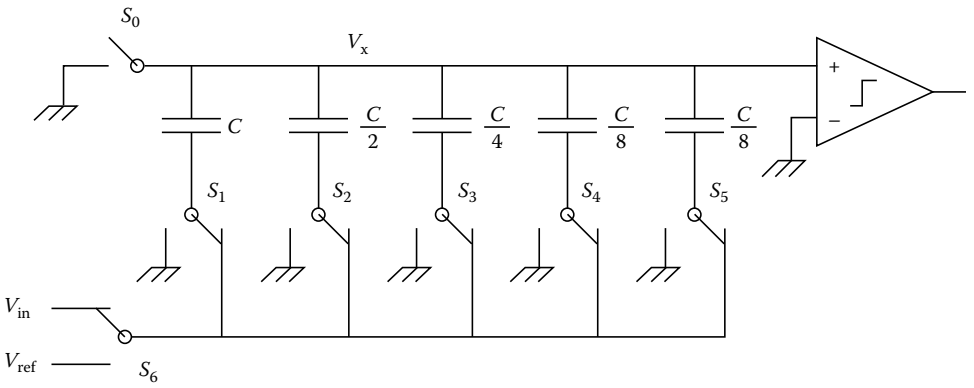


FIGURE 10.49 Charge-distribution implementation of the successive approximation architecture.

If $V_x > 0$ then the comparator output goes high, signifying that $V_{in} < (V_{ref}/2)$ and switch S_1 is switched back to ground. If the comparator output is low, then $V_{in} > (V_{ref}/2)$ and the switch S_1 is left connected to V_{ref} and the MSB is set high. In a similar fashion the next bit, MSB-1, is evaluated. This procedure is continued until all N-bits have resolved. After the conversion process the voltage at the top plate is such that

$$V_x = -V_{in} + \left\{ b_3 \frac{V_{ref}}{2^1} + b_2 \frac{V_{ref}}{2^2} + b_1 \frac{V_{ref}}{2^3} + b_0 \frac{V_{ref}}{2^0} \right\} \tag{10.23a}$$

$$-1 \text{ LSB} < V_x < 0 \tag{10.23b}$$

where b_i is {0, 1} depending upon if bit$_i$ was set to 0 or 1.

One of the advantages of the charge-redistribution topology is that the parasitic capacitance from the switches has little effect on the accuracy. Additionally, the clock feedthrough from switch S_0 only causes an offset and the clock feedthrough from switches $\{S_1, S_2, \ldots, S_5\}$ is input signal independent because they are always connected to either ground or V_{ref}. However, any mismatch in the binary ratios of the capacitors in the array causes nonlinearity, which limits the accuracy to 10 or 12 bits.

Self-calibration successive approximation converters. Fortunately, self-calibrating [7] techniques have been introduced that correct for errors in the binary ratios of the capacitors. Figure 10.50 is the block diagram for a successive approximation-based self-calibrating ADC. The circuit consists of an N-bit binary weighted capacitor array main DAC, an M-bit resistor string sub-DAC, and a calibration DAC. Digital logic is used to control the circuit during calibration and also to store the error voltages.

Let each weighted capacitor C_i have a normalized error in its ratio $(1 + \varepsilon_i)$ from its ideal value:

$$C_i = 2^{i-1} C(1 + \varepsilon_i) \tag{10.24}$$

Each capacitor contributes an error voltage at the top plate which is equal to

$$V_{e_i} = \frac{V_{ref}}{2^N} 2^{i-1} \varepsilon_i \quad i = 1B, 2, \ldots, N \tag{10.25}$$

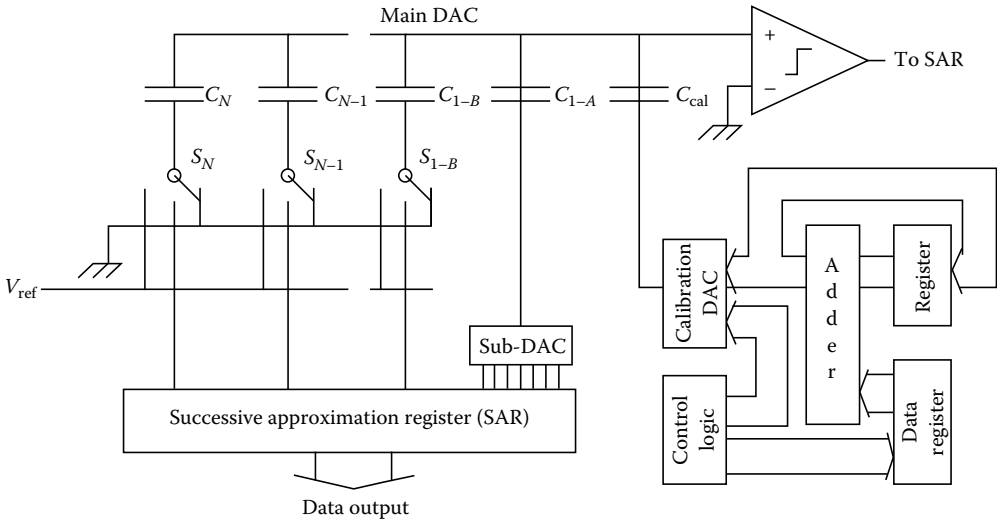


FIGURE 10.50 Self-calibration charge redistribution converter.

Therefore, the total linearity error is equal to

$$V_{\text{error}} = \sum_{i=1B}^{N} V_{\varepsilon_i} b_i \tag{10.26}$$

where b_i is the logic value of the ith bit.

The calibration cycle begins by measuring the error contribution from the largest capacitor and progressing to the smallest. The error from the MSB capacitor is evaluated by closing S_0 and setting switches $\{S_1, S_2, \ldots, S_5\}$ such that all the capacitors except C_{MSB} are charged to V_{ref}. Next, the switch S_0 is opened and switches $\{S_1, S_2, \ldots, S_5\}$ are switched to connect the bottom plates to ground. Under ideal conditions, i.e., $C_{\text{MSB}} = 2^{N-1}C$, the voltage at the top plate is equal to zero. It should be noted that the total capacitance is equal to $2C$. However, because $C_{\text{MSB}} = 2^{N-1}C(1 + \varepsilon_{\text{MSB}})$, the top plate voltage $V_x = (V_{\text{ref}}/2)\varepsilon_{\text{MSB}}$, such that $V_{x_{\text{MSB}}} = 2V_{\varepsilon_{\text{MSB}}}$ Therefore, the error voltage at the top plate is a direct measure of the corresponding error in the capacitor ratio. A successive approximation search using the sub-DAC is used to measure these voltages. The relationship between the measured residual voltage and the error voltage is equal to

$$V_{\varepsilon_i} = \frac{1}{2}\left\{ V_{x_i} - \sum_{j=i+1}^{N} V_{\varepsilon_i} \right\} \tag{10.27}$$

which corresponds to the equivalent error terms on the digital side. These digital correction terms are stored and subsequently added or subtracted during the normal operation cycle. Self-calibration improves the resolution of successive approximation converters to approximately 15 or 16 bits.

Algorithmic converters. As stated earlier, the algorithmic ADC is formed by modifying a pipelined converter. Here, the pipeline has been closed to form a loop. All N-bits are evaluated by a single stage, therefore implying that a N-bit conversion requires N clock cycles. A block diagram for the algorithmic converter is shown in Figure 10.51 [6] and consists of a S/H, a $2\times$ amplifier, a comparator, and a reference subtraction circuit. The circuit operates as follows. The input is first sampled and held by

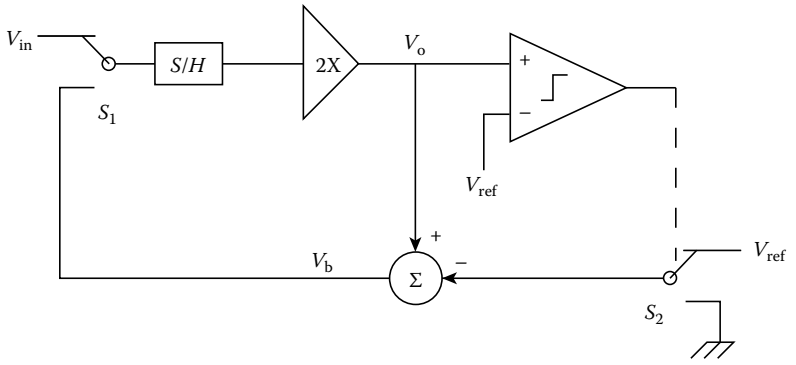


FIGURE 10.51 Algorithmic ADC block diagram.

setting S_1 to V_{in}. This signal is then multiplied by 2 (by the 2× amplifier). The result of this multiplication, V_o, is compared to V_{ref}. If $V_{o_N} > V_{ref}$ then the MSB, b_N, is set to 1 or it is set to 0. If b_N is equal to 1, then S_2 is connected to V_{ref} such that V_{b_N} is equal to

$$V_{b_N} = 2V_{0_N} - b_N V_{ref} \quad b_N = \{0, 1\} \tag{10.28}$$

This voltage is then sample-and-held and used to evaluate the MSB-1 bit. This procedure continues until all N-bits are resolved. The general expression for V_o is equal to

$$V_{0_i} = [2V_{0_{i-1}} - b_i V_{ref}]z^{-1} \tag{10.29}$$

where b_i is the comparator output for the ith evaluation and z^{-1} implies a delay of one clock period.

A circuit implementation for this ADC topology is shown in Figure 10.52 [10]. This circuit uses three amplifiers, five ratio-matched capacitors (C_1 to C_5), an arbitrary valued capacitor, C_6, and a comparator. Two implifiers and the capacitors (C_1 to C_5) form the recirculating register and the gain of two amplifiers. The amplifier A_3 and capacitor C_6 form an offset compensated comparator. The switches controlled by V_3, V_4, and V_5 load the input or selectively subtract the reference voltage. The conversion is started by setting V_1, V_2, and V_3 high. This forces V_x and V_y to 0 and loads V_{in} into C_1. Then, V_1 is set low and V_5 is set high. Therefore, the charge $V_{in} * C_1$ is transferred from C_1 to C_2. C_1 is made to be equal to C_2, therefore, $V_x = V_{in}$(C_3 is also charged to V_x). Because V_1 has been set low the comparator output goes high if $V_{in} > 0$, or else it remains low. This determines the MSB. The MSB-1 is determined by setting V_2 low and setting V_1 high. This forces the charge from C_3 to transfer to C_4($V_4 = V_x$; C_5 is also charged to V_4). During the same period C_1 is connected to ground if MSB = 1, or it is connected to V_{ref}. Next, V_2 is set low and V_1 is set high, while C_1 is switched from ground to V_{ref} for V_{ref} to ground. This transfers a charge equivalent to $C_1 \pm V_{ref}$ from C_1 to C_2 and transfers the charge in C_5, $C_5 * V_y$, to C_2. The capacitor C_5 is made to be twice as large as C_2, therefore, the voltage at V_x is equal to $2 * V_{in} \pm V_{ref}$. This process is repeated and the comparator determines bit MSB-1. This circuit has been shown to provide up to 10 bits of resolution at a maximum conversion rate of 200 kHz.

The maximum resolution of the algorithmic converter is limited by the ratio matching of the capacitors, clock feedthrough, capacitor voltage coefficient, parasitic capacitance, and offset voltages. The previous topology solves the problem of parasitic capacitances and amplifier offset voltage, however, its maximum resolution is limited by the ratio matching of the capacitors that are used to realize the gain of two amplifiers. This problem is partially resolved by using a ratio-independent multiply-by-two algorithm [6] to increase the maximum resolution to the 12 bit level. The ratio-independent multiply-by-two algorithm is easily explained by the circuit shown in Figure 10.53. During ϕ_1, capacitor C_1 is

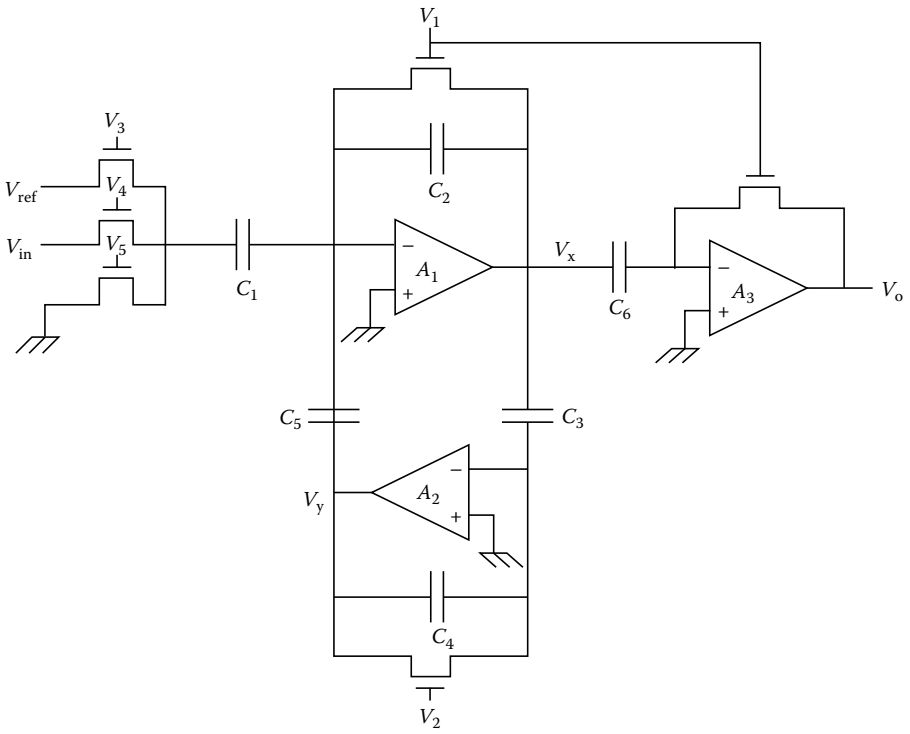


FIGURE 10.52 Example circuit implementation of the algorithmic converter.

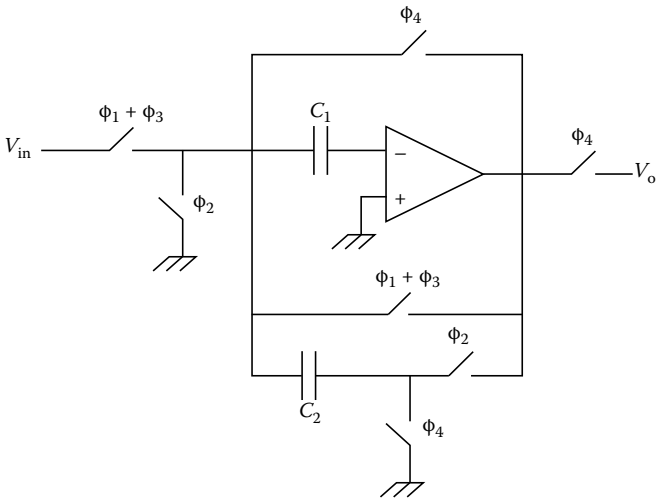


FIGURE 10.53 Ratio-independent multiply-by-two circuit.

charged to V_{in}. This charge is then transferred onto C_2 during ϕ_2. The charge on C_2 is equal to $V_{in} * C_1$. During ϕ_3 C_2 is disconnected from the feedback path and V_{in} is once again sampled onto C_1. During ϕ_4 the charge in C_2 is added to C_1. The total charge in C_1 is now equal to $C_1\ V_{in} + C_1\ V_{in} = 2C_1V_{in}$ and is completely independent of the value of C_2. Therefore, the voltage at the output at ϕ_4 is equal to $2V_{in}$. The only constraint is that the input voltage be held steady, i.e., S/H during ϕ_1 and ϕ_3.

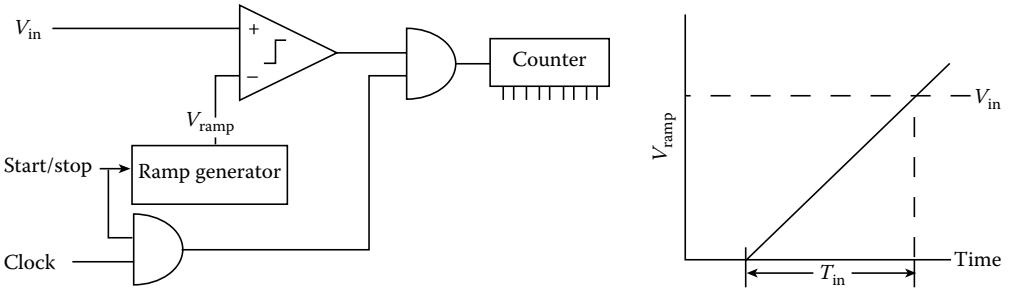


FIGURE 10.54 Single-slope integrating converter.

10.2.2.3 2^N-Clock Converters

The basic principle of the integrating converter can be explained with the help of Figure 10.54. A comparator compares the input signal with the output of a ramp voltage generator. The ramp voltage generator is zeroed after each measurement. The output of this comparator is used to gate the clock to an interval counter. The counter output corresponding to the ramp time T_{in} provides an accurate measure of the input voltage. The input voltage V_{in} is equal to $T_{in} * U$, where U is the ramp rate. Because the absolute values of components are not well controlled and also because of the large offset voltages associated with MOS amplifiers and comparators, a calibration or a reference cycle is usually added to calculate the ramp rate and the offset voltage. A simple circuit for a single-slope integrating converter that includes the calibration cycle is shown in Figure 10.55 [4].

The ramp voltage is generated using a constant current source to charge a capacitor. The ramp voltage, V_{ramp}, is equal to $\int_0^t (I/c)dt$, which is equal to $(I\Delta t)/c$ for a constant current I. The ramp voltage is compared against the analog ground voltages, V_{in} and V_{ref}, respectively. The addition of the third calibration cycle eliminates any offset errors. The final resolution is dependent only on the linearity of the ramp generator, i.e., the linearity of the current source. In the single-slope approach just described, the calibration is done in digital. However, the complete calibration can be performed in analog as well, as in the dual-slope approach. Further improvements include a charge balancing technique [4] that

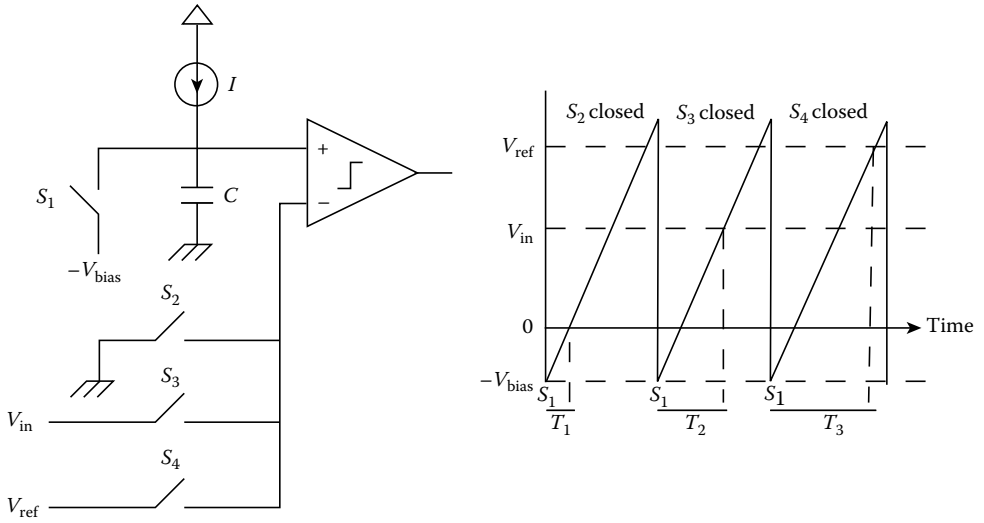


FIGURE 10.55 Single-slope integrating converter with calibration cycle.

uses an oscillating integration process to keep the voltage across the capacitor closer to zero, thereby reducing the linearity constraints on the ramp generator. The primary advantage of the integrating converter is the small number of precision analog components that are required to generate extremely high resolution. However, the primary disadvantage is the conversion time required. It takes 2^N clock cycles to generate a N-bit conversion.

10.2.3 Oversampled Converters

Oversampling converters have the advantage over Nyquist rate converters in that they do not require very tight tolerances from the analog components and also because they simplify the design of the anti-alias filter. Examples of oversampling converters include the noise-shaping architecture and the interpolative architecture. Our discussion centers around noise-shaping converters.

If the analog input signal V_{in} has a frequency spectrum from 0 to f_0 then $2f_0$ is defined as the Nyquist rate. Oversampling converters sample the input at a rate larger than the Nyquist frequency. If f_s is the sampling rate, then $(f_s)/(2f_0) = \text{OSR}$ is called the oversampling ratio. Oversampling converters use "signal averaging" along with a low-resolution converter to provide extremely high resolution. This technique can best be understood by considering the following example in Figure 10.56.

Let the input be exactly in the middle of V_n and V_{n+1} and let it be sampled a number of times. If, in addition to the input signal, we add some random noise, then for a large number of samples the output would fall on V_n 50% of the time and on V_{n+1} the other 50% of the time. If the signal was a littler closer to V_{n+1}, then the percentage of times the output falls on V_{n+1} would increase. Using this averaging technique we can get a better estimate of the input signal. However, in simple oversampling the resolution only increases by $\sqrt{n}$, where n is the number of samples of V_{in} that are averaged. Therefore, to increase the resolution of the converter by one additional bit we are required to increase the number of samples by 4×.

Noise-shaping converters use feedback to generate the necessary noise and additionally perform frequency shaping of the noise spectrum to reduce the amount of oversampling necessary. This can be illustrated with the help of Figure 10.57. The output from H_1 is quantized by an N-bit ADC. This digital value is then converted to an analog value by the N-bit DAC. This value is subtracted from the input and the result is sent to H_1. Here, we assume an N-bit converter for simplicity, however, for the special case in which $N = 1$ the noise-shaping converter is called a sigma-delta converter. The quantization process approximates an analog value by a finite-resolution digital value. This step introduces a quantization error, Q_n. Further, if we assume that the quantization error is not correlated to the input, the system can now be modeled as a linear system, as shown in Figure 10.58. Here, we note that the error introduced by the analog-to-digital process is modeled by Q_n. The output voltage for this system can now be written as

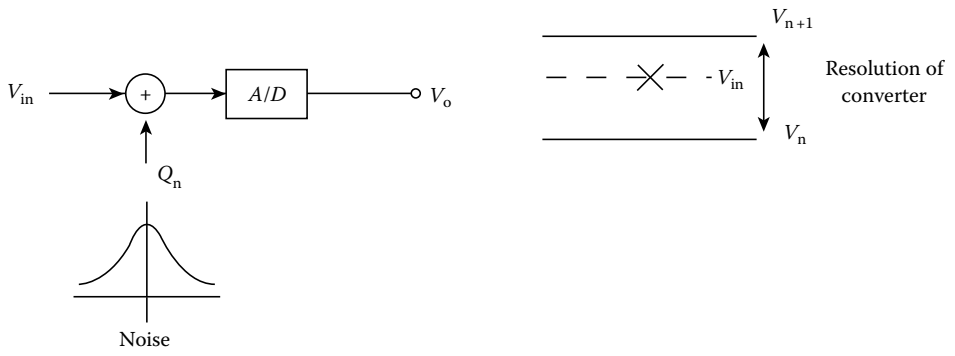


FIGURE 10.56 Higher resolution provided by oversampling.

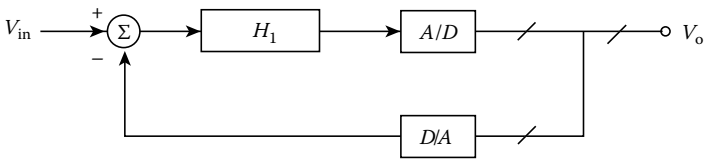


FIGURE 10.57 Noise-shaping oversampling converters.

$$V_0 = \frac{Q_n}{[1+H_1]} + \frac{V_{in}H_1}{[1+H_1]} \tag{10.30}$$

Data converters are sampled data systems, and as such are easier to analyze in the Z-domain. For most sigma-delta converters H_1 has the characteristics of a low-pass filter and is usually implemented as a switched-capacitor integrator. MOS switched-capacitor integrators can be implemented with either a delay in the forward signal path or a delay in the feedback path, and can be modeled in the Z-domain by Figures 10.59 and 10.60, respectively.

We use the first integrator architecture because it simplifies some of the algebra. For a first-order sigma-delta converter H_1 is realized as a simple switched-capacitor integrator, i.e., $H_1 = (z^{-1})/(1-z^{-1})$. Therefore, Figure 10.58 can now be drawn as Figure 10.61. Replacing H_1 by $(z^{-1})/(1-z^{-1})$ in Equation 10.30 we can write the transfer function for the first-order sigma-delta converter as

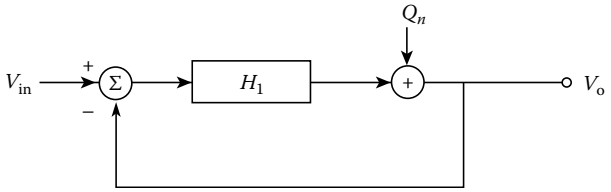


FIGURE 10.58 Linear system model of noise-shaping converter.

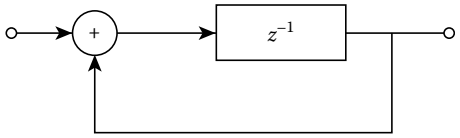


FIGURE 10.59 Forward path delay integrator.

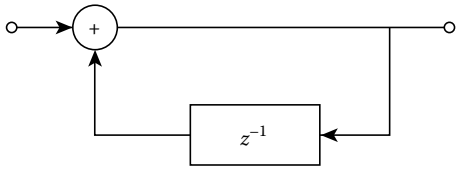


FIGURE 10.60 Feedback path delay integrator.

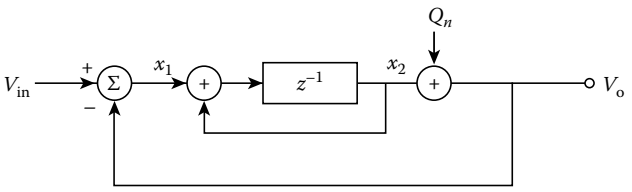


FIGURE 10.61 First-order noise-shaping converter.

$$V_o = V_{in}z^{-1} + Q_n(1 - z^{-1}) \tag{10.31}$$

As can be seen from Equation 10.31, the output is a delayed version of the input plus the quantization noise multiplied by the factor $(1 - z^{-1})$. This function has a high-pass characteristic, as shown in Figure 10.62. We note here that the quantization noise is substantially reduced at lower frequencies and increases slightly at higher frequencies. In this figure, f_o is the input signal bandwidth and $f_s/2 = \pi$ corresponds to the Nyquist rate of the oversampling converter. For simplicity the quantization noise is usually assumed to be white* with a spectral density equal to $e_{rms}\sqrt{2/f_s}$. Therefore, the magnitude of the output noise spectrum can be written as

$$N(f) = e_{rms}\sqrt{\frac{2}{f_s}}\left|1 - z^{-1}\right| = 2e_{rms}\sqrt{\frac{2}{f_s}}\sin\left(\frac{\pi f}{f_s}\right) \tag{10.32}$$

Further, if $f_o \ll f_s$ we can approximate the root mean square noise in the signal band, $(0 < f < f_o)$, by

$$N_{f_0} \approx e_{rms}\frac{\pi}{3}\left(\frac{2f_0}{f_s}\right)^{3/2} \tag{10.33}$$

As the OSR increases the quantization noise in the signal band decreases; i.e., for a doubling of the OSR the quantization noise drops by 20 $\log(2)^{3/2} \approx 9$ dB. Therefore, for each doubling of the OSR we effectively increase the resolution of the oversampling converter by an additional 1.5 bits.

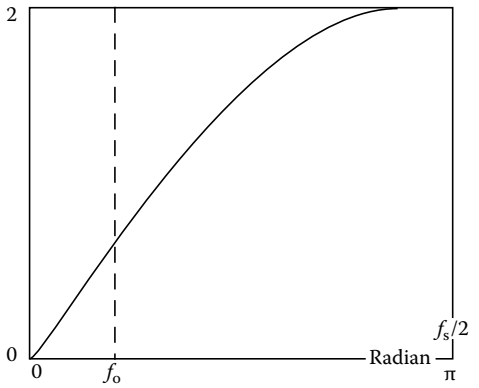


FIGURE 10.62 Magnitude response of the function $(1 - z^{-1})$.

* Quantization noise is clearly not uncorrelated or white for the first-order sigma-delta modulator, but becomes increasingly so for the higher-order systems.

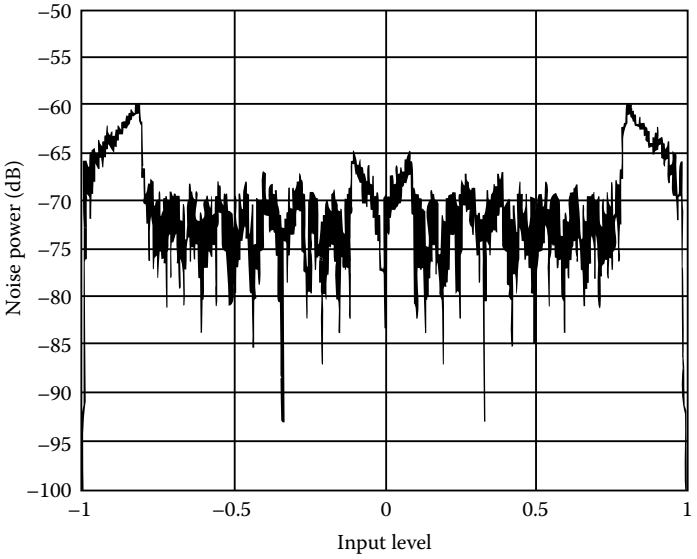


FIGURE 10.63 Pattern noise for a first-order sigma-delta modulator for DC inputs.

The previous analysis was based on the assumption that the quantization noise was not correlated to the input and uniformly distributed across the Nyquist band. We now reexamine these assumptions. The assumption that the quantization noise is not correlated with the input only holds for extremely busy input signals. This is particularly not true for the first-order modulator assumed in the analysis above, such that for extremely low frequency or DC inputs the first-order modulator generates pattern noise (also called tones), as shown in Figure 10.63. The peaks of the pattern noise occur at input voltages that are integer divisors of the quantization step. It is possible to provide a conceptual explanation for this occurrence. For example, for an input that is an integer divisor of the quantization level the digital output of the quantizer repeats itself at an extremely low frequency. This low-frequency repetition causes noise power to be introduced into the signal band. The quantization noise for second- and higher-order models is significantly more uncorrelated and is usually assumed to be white.

The quantization error has a value that is limited to $+\frac{1}{2}$ LSB of the quantizer (the ADC in Figure 10.47). If we assume that the quantization noise is white and uniformly distributed over the quantization level, then the average noise quantization is equal to

$$\int_{-\frac{1}{2}\mathrm{LSB}}^{+\frac{1}{2}\mathrm{LSB}} x^2 \mathrm{d}x = \frac{\mathrm{LSB}^2}{12} = P_n \tag{10.34}$$

Because the quantization noise is sampled at the clock frequency f_s, the entire noise power is aliased back into the overall converter Nyquist ban $[0-(f_s/2)]$. Therefore, the spectral density of the quantization noise is equal to

$$P_n = \frac{\mathrm{LSB}^2}{12} = \int_0^{f_s/2} n_e(f)^2 \delta f = n_e(f)^2 \frac{f_s}{2} \tag{10.35a}$$

$$n_e(f) = P_n\sqrt{\frac{2}{f_s}} \tag{10.35b}$$

The SNR for an ADC is defined as $10\log(P_s/P_n)$, where P_s is the signal power. The signal power is highly waveform dependent. For example, the P_s for a full-scale sine wave input, $(A/2)\sin(\omega T)$, which is applied to an N-bit quantizer, can be written in terms of the quantization level as

$$\frac{A^2}{8} = \frac{[(2^N - 1)\text{LSB}]^2}{8} \tag{10.36}$$

Therefore,

$$\text{SNR} = 10\log\left(\frac{P_s}{P_n}\right) = 10\log\left[\frac{12(2^N - 1)^2}{8}\right] \tag{10.37}$$

10.2.3.1 Higher-Order Modulators

In Figure 10.61, we replaced H_1 for Figure 10.58 with a first-order integrator. Clearly, H_1 can be replaced by other higher-order functions that have a low-pass characteristic.* For example, in Figure 10.64 we show a second-order modulator. This modulator uses one forward delay integrator and one feedback delay integrator to avoid stability problems. The output voltage for this figure can be written as

$$V_0 = V_{\text{in}}z^{-1} + Q_n(1 - z^{-1})^2 \tag{10.38}$$

Note that the quantization noise is shaped by the second-order difference equation. This serves to further reduce the quantization noise at low frequencies. However, a further increase in the noise occurs at higher frequencies. A comparison of the noise shaping offered by the first and second-order modulators is shown in Figure 10.65. Once again, assuming that $f_o \ll f_s$ we can write an expression for the root mean square noise in the signal band for the second-order modulator as

$$N_{f_0} \approx e_{\text{rms}}\frac{\pi^2}{\sqrt{5}}\left(\frac{2f_0}{f_s}\right)^{5/2} \tag{10.39}$$

The noise power in the signal bandwidth falls by 15 dB for every doubling of the OSR. One of the added advantages of the second-order modulator over the first-order modulator is that quantization noise has been shown to be less correlated to the input, therefore, less pattern noise.

From our analysis so far, it would seem that increasing the order of the filter would reduce the necessary OSR for a given resolution. This is true, however, the simple Candy-style modulator (shown in

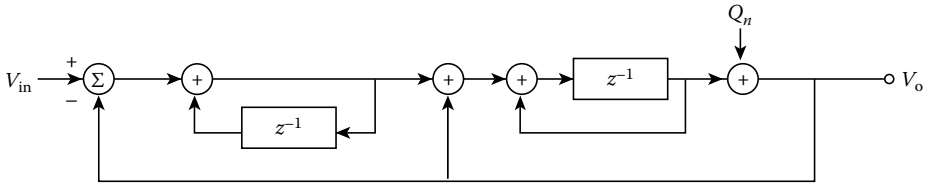


FIGURE 10.64 Second-order modulator block diagram.

* Actually, it is not necessary that they have low-pass characteristics. Bandpass characteristics may be preferred if the input signal is to be bandlimited.

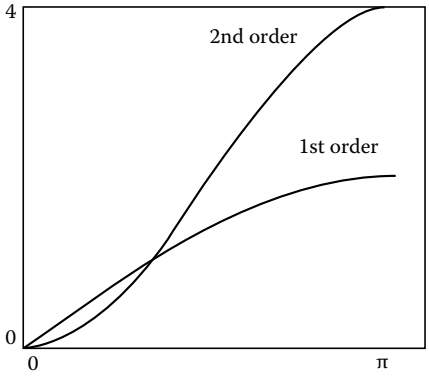


FIGURE 10.65 Noise shaping due to the second-order modulator in Figure 10.62.

Figures 10.61 and 10.64) with orders >2 results in stability problems. This is because for higher-order modulators the later integrator stages are easily overloaded and saturated. This in turn increases the noise in the signal band. However, higher-order modulators can be realized by using a cascade of lower order modulators in the MASH architecture [8]. In the cascaded MASH technique, both the digital output and the output of the integrator of each lower-order modulator is passed on to the next module. A second-order MASH architecture using two cascaded first-order sections is shown in Figure 10.66. It can be shown that the output is equal to

$$Y = z^{-2}X - Q_{n_2}(1 - z^{-1})^2 \tag{10.40}$$

Once again, we note that the quantization noise is multiplied by the second-order difference equation. The sign in front of the noise term is not important. However, for complete cancellation of the quantization noise from the first integrator the gain of the first loop needs to be identical to the gain of the second loop. Therefore, the amplifier gain and capacitor matching become extremely important. It has been shown that a 1% matching and an op-amp gain of 80 dB are sufficient for 16 bits of accuracy [8].

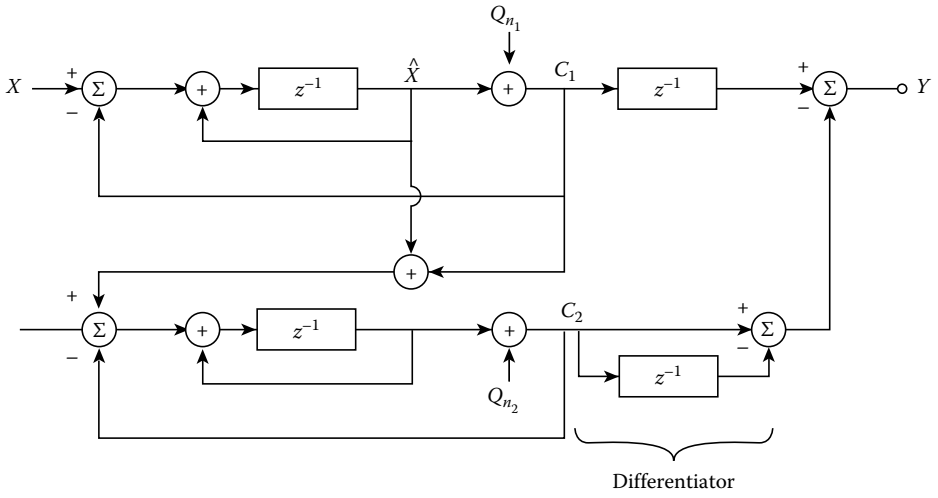


FIGURE 10.66 MASH architecture for a second-order modulator.

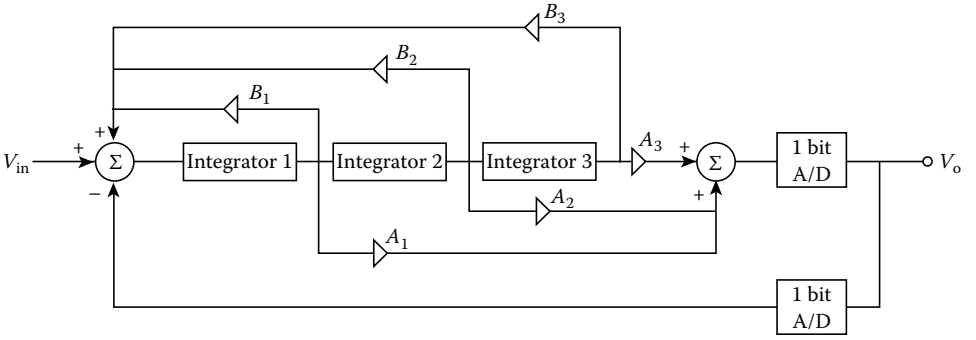


FIGURE 10.67 Finite pole-zero loop filter higher-order modulator.

An alternate methodology to stabilize higher-order oversampled coders is the use of finite poles and zeroes for the loop filter [16], H_1 in Figure 10.57. Up until now, all the loop filters have been integrators with poles at DC and zeroes at extremely high frequencies. The loop filter can be realized using additional feedback and feedforward paths as shown in Figure 10.67. A third-order modulator is shown in this figure. Having finite poles and zeroes serves two purposes: (1) the nonzero poles function to reduce the in-band noise by flattering the quantization noise transfer function at low frequencies, (2) the finite zeroes function to reduce the magnitude of the quantization noise at high frequencies. By reducing the magnitude of the quantization noise at high frequencies, even higher-order modulators can be made stable. Additionally, these modulators have been shown to be devoid of pattern noise artifacts.

10.2.3.2 Multibit Quantizers

The primary reason for using single-bit or two-level quantizers is their inherent perfect linearity. Because only two levels can exist, a straight line can always be drawn between these two levels. On the other hand number of advantages are found in using multibit quantizers in oversampling converters. The quantization noise generated in the multibit-based noise-shaping converter is significantly more "white" and uncorrelated with the input signal, thereby reducing the probability of pattern noise. Additionally, the quantization noise power goes down exponentially as the number of bits in the quantizer increases. However, the primary problem associated with multilevel quantizers is the nonlinearity errors present with the DAC in the modulator loop. This problem can be illustrated with the help of Figure 10.68.

In Figure 10.68, the error resulting from the nonlinearity in the multibit ADC is included as AD_{NL} and the error resulting from the nonlinearity in the multibit DAC is included as DA_{NL}. The output voltage is given by Equation 10.43. Here, note that the analog-to-digital nonlinearity is suppressed by the loop filter, while the digital-to-analog nonlinearity is only subjected to a unit delay. Therefore, any

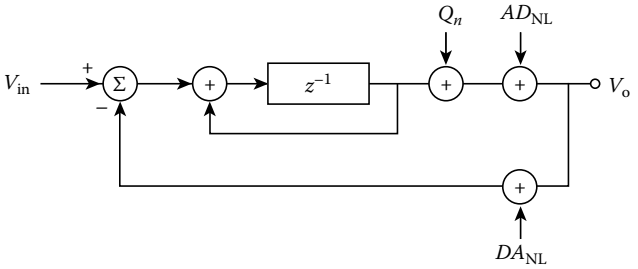


FIGURE 10.68 Model for nonlinearity associated with multibit quantizers.

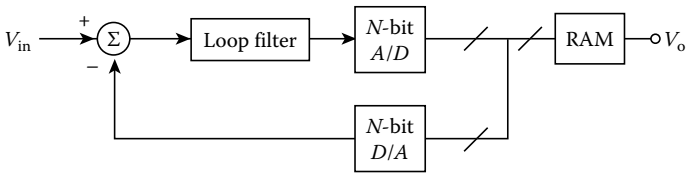


FIGURE 10.69 Digital error correction for multibit quantizers.

digital-to-analog nonlinearity directly appears in the output. A number of methods have been applied to reduce the effects of nonlinearity associated with multibit quantizers. The two most promising methods that have emerged are digital error correction [18] and dynamic element matching [17,19].

$$V_0 = V_{in}z^{-1} - DA_{NL}z^{-1} + Q_n(1 - z^{-1}) + AD_{NL}(1 - z^{-1}) \tag{10.41}$$

A block diagram for digital error correction for multibit quantizer-based noise-shaping converters is shown in Figure 10.69. The random access memory (RAM) and the multibit DAC have the same input signal. Because of the high gain in the loop at low frequencies the output of the DAC is almost identical to the input voltage, V_{in}. Now, if the digital RAM is programmed to generate the exact digital equivalent of the digital-to-analog output for any digital input, then the RAM output and the digital-to-analog output will be identical to each other. Because the output of the DAC is almost identical to the input voltage, the output voltage will also be the exact digital equivalent of the analog input. The RAM can be programmed by reconfiguring the modulator stages and feeding the system with a multibit digital ramp [18].

In the dynamic element matching approach, the various analog elements that are used to generate the different analog voltage levels are dynamically swapped around. The various elements can be swapped randomly [17] or in a periodic fashion [19]. The use of random permutations translates the nonlinearity of the DAC into random noise that is distributed throughout the oversampling converter Nyquist range. This method virtually eliminates errors due to nonlinearity, but unfortunately it also increases the noise level in the signal band. In a variation of this basic technique, the various analog elements are swapped in a periodic fashion such that the nonlinearity in the DAC is translated into noise at higher frequencies. Individual level averaging further eliminates the possibility of pattern noise within the signal band [19].

10.2.3.3 Technology Constraints

One of the primary reasons for using sigma-delta converters is that they do not require good matching among the analog components. Therefore, for the two-level sigma-delta converter the nonidealities are introduced primarily by the integrator loop. To aid in the analysis of the various technology constraints we shall consider a particular implementation of an integrator (Figure 10.70). The ideal transfer function for this circuit is given by $-1/(1 - z^{-1})$. To realize this ideal transfer function the circuit relies on the virtual ground generated at the negative input of the integrator to accomplish complete charge transfer during each clock period. However, limited amplifier gain does not generate a perfect virtual ground,

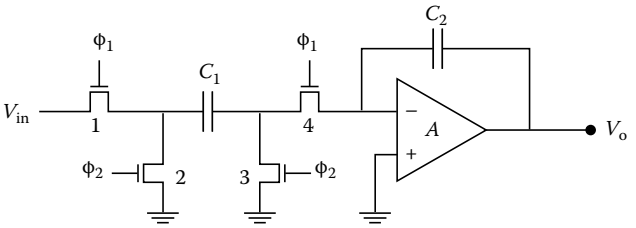


FIGURE 10.70 Example circuit implementation for a switched-capacitor integrator.

thereby not accomplishing the complete transfer of charge during each clock period. The effect of the limited gain is similar to a leaky integrator and the transfer function for the leaky integrator can be written as

$$H(z) = \frac{-1}{1 - \alpha z^{-1}} \tag{10.42a}$$

where

$$\alpha = \frac{1}{1 - \frac{1}{A}\left(1 + \frac{C_1}{C_2}\right)} \approx \frac{1}{1 - \frac{2}{A}} \tag{10.42b}$$

The net effect of finite gain is to increase the modulation noise in the signal band as illustrated by Figure 10.71 for the first-order modulator. In this figure, the *X*-axis is plotted from 0 to 1 rather than the complete Nyquist band to emphasize the signal band. The noise transfer function has been plotted for a number of amplifier gains. When compared to Figure 10.62, note the increase in the noise level in the signal band. The effect of finite gain is felt throughout the input signal magnitude range as shown in Figure 10.72. The graph for Figure 10.71 was generated using the linearized model for the modulator

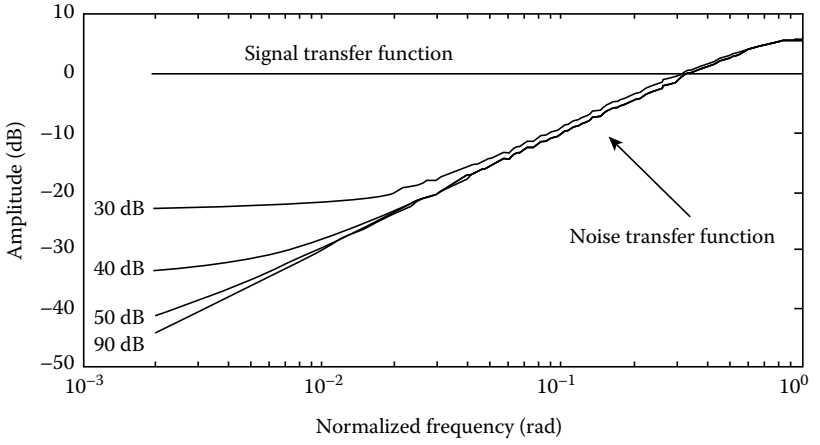


FIGURE 10.71 Effect of finite amplifier gain on noise transfer function using a linear model.

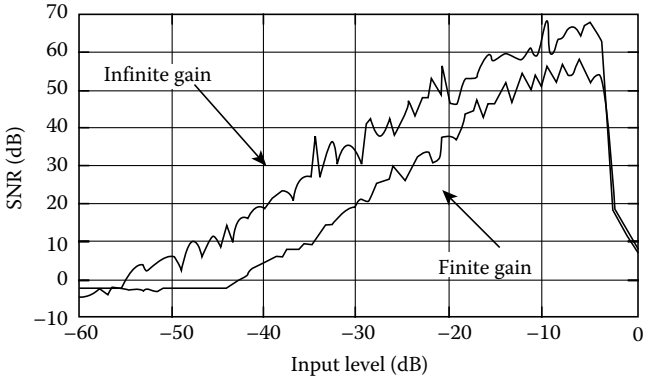


FIGURE 10.72 Effect of finite amplifier gain using difference equation simulations.

presented in Equation 10.21 and the graph in Figure 10.72 was generated using the difference equation method. The difference equation method does not make any assumptions about linearity nor does it assume that the input is uncorrelated with the quantization noise, however, it requires considerably more simulation time. Because of oversampling the bandwidth requirements for the op-amps in the integrators are usually large. Unfortunately, it is extremely difficult to realize extremely high gain and extremely high bandwidth amplifiers in MOS. One solution that attempts to mitigate the finite gain effect is to estimate the amount of incomplete charge transfer and compensate for it [5].

Circuit noise provides additional limitations to the maximum resolution realizable by an oversampled converter. The primary noise sources are the thermal noise generated by the switches in the integrator, amplifier noise, charge injection, and clock feedthrough from the switches. Because of sampling, the thermal noise associated with the finite on resistance of the switches is aliased back into the Nyquist band of the oversampling converter. The total noise aliased into the baseband for large bandwidth amplifiers is equal to kT/C for each switch pair, where k is the Boltzmann constant, T is the temperature in degrees Kelvin, and C is the value of the sampling capacitor in the integrator. For the parasitic insensitive integrator in Figure 10.70, the total noise from this source is equal to $2kT/C$. This noise is evenly spread across the Nyquist band, but only the fraction $2f_o/f_s$ of this noise appears in the signal band. The rest is filtered out by the digital LPF. Using this constraint, for a full-scale sine wave input the minimum sampling capacitance is given by

$$C_{\min} = 16 \cdot k\mathrm{T} \cdot \mathrm{SNR}_{\mathrm{desired}} \tag{10.43}$$

The inband portion of the amplifier noise is also added to the output signal. In general, only the noise of the first amplifier is important for higher-order converters. For MOS amplifiers the flicker noise component is significantly more important as it tends to dominate in the signal band. When necessary, correlated double sampling techniques [5] can be used to reduce the effect of this noise source. Correlated double sampling, or autozeroing as it is sometimes called, has the added benefit that it eliminates any amplifier offset voltages. It is usually important to remove this offset voltage only for data acquisition applications.

Because tight component matching is not required of the analog components, sigma-delta converters are particularly well suited for mixed-signal applications. However, having digital circuits on the sample chip increases the switching noise that is injected into the substrate and into the power supply lines. Any portion of this noise that lies in the signal band is added to the input signal. Therefore, fully differential integrator topologies should be used for high-resolution converters. Substrate and supply noise are common-mode signals and are reduced by the common-mode rejection ratio of the amplifier when using fully differential circuits.

In addition to the amplifier and the switching noise the charge injection from switches also sets a limit on the maximum resolution attainable component. Charge injection from switches has a signal-dependent component and a signal-independent component. The effect of the signal-independent component is to introduce an additional offset error that can easily be calibrated out, if necessary. However, the signal-dependent component, particularly from the input sampling transistor (transistor M_1 in Figure 10.73), cannot be distinguished from input signal. This signal-dependent component is highly nonlinear and can be reduced substantially by using proper clock phasing. Signal-dependent charge injection from transistors M_1 and M_2 in Figure 10.73 can be canceled to first order by delaying the turn off of ϕ_1' slightly [6].

A number of topologies for the digital low-pass filters have been tried. However, it has been shown that simple finite impulse response (sinc) filters are probably the optimal choice. It has been shown that the number of stages of sinc filtering necessary is equal to the modulator order plus 1 [2]. Noise-shaping converters have the ability to provide extremely high resolution. However, care must be used when using simple linear assumptions. Clearly, for the first-order modulator the white noise assumption

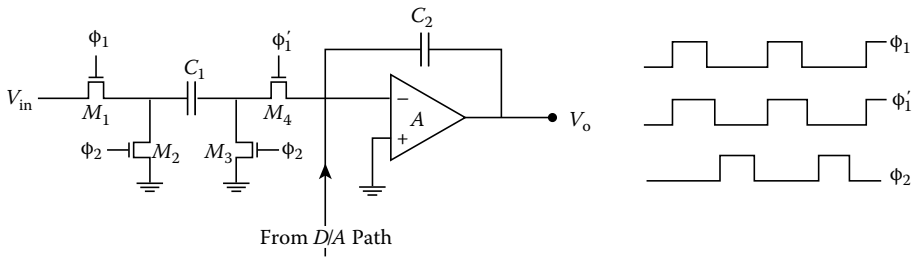


FIGURE 10.73 Proper clock phase to eliminate signal-dependent clock feedthrough.

breaks down. Additionally, it has been shown that the simple linear model overestimates the realizable SNR. For example, the linearized model overestimates the attainable SNR by as much as 14 dB for the second-order modulator.

Acknowledgment

The author acknowledges the help of his students in completing this manuscript, particularly Feng Wang.

References

1. W. C. Black, High-speed CMOS A/D conversion techniques, PhD thesis, Berkeley: University of California, 1980.
2. J. C. Candy and G. C. Temes, *Oversampling Methods for A/D and D/A Conversion*, New York: IEEE Press, 1992.
3. M. J. Demler, *High-Speed Analog-to-Digital Conversion*, New York: Academic Press, 1991.
4. P. R. Gray and D. A. Hodges, All-MOS analog-digital conversion techniques. *IEEE Trans. Circuits Syst*., CAS-25(7), 482–489, 1978.
5. P. J. Hurst and R. A. Levinson, Delta-sigma A/Ds with reduced sensitivity to op-amp noise and gain. *Proceedings of IEEE Int. Symp. Circuits Syst.*, Portland, Oregon, pp. 254–257, 1989.
6. P. W. Li, M. J. Chin, P. R. Gray, and R. Castello, A ratio-independent algorithmic analog-to-digital conversion technique. *IEEE J. Solid-State Circuits*, SC-19(6), 828–836, 1984.
7. H. S. Lee, D. A. Hodges, and P. R. Gray, A self-calibrating 15 bit CMOS A/D converter. *IEEE J. Solid-State Circuits*, SC-19(6), 813–819, 1983.
8. Y. M. Matsuya, K. Uchimura, and A. Iwata, A 16-bit oversampling A-to-D conversion technology using triple integration noise shaping. *IEEE J. Solid-State Circuits*, SC-22, 921–929, 1987.
9. J. L. McCreary and P. R. Gray, All-MOS charge redistribution analog-to-digital conversion techniques I. *IEEE J. Solid-State Circuits*, SC-10, 371–379, 1975.
10. R. H. McCharles, V. A. Saletore, W. C. Black, and D. A. Hodges, An algorithmic analog-to-digital converter. *Proceedings of the IEEE Int. Solid-State Circuits Conf*., San Francisco, California, 1977.
11. S. Masuda, Y. Kitamura, S. Ohya, and M. Kikuchi, A CMOS pipelined algorithmic A/D converter, *Proceedings of the IEEE Custom Integrated Circuits Conf*., San Jose, California, pp. 559–562, 1984.
12. G. C. Temes, F. J. Wang, and K. Watanabe, Novel pipeline data converters. *Proceedings of the IEEE Int. Symp. Circuits Systems*, Helsinki, Finland, pp. 1943–1946, 1988.
13. R. Unbehauen and A. Cichocki, *MOS Switched-Capacitor and Continuous-Time Integrated Circuits and Systems*, New York: Springer-Verlag, 1989.
14. S. H. Lewis, Video-rate analog-to-digital conversion using pipelined architectures, PhD thesis, Berkeley, CA: University of California, 1987.

15. S. Sutarja and P. R. Gray, A pipelined 13-bit, 250-ks/s, 5-V analog-to-digital converter. *IEEE J. Solid-State Circuits*, SC-23, 1316–1323, 1988.
16. K. Chao, S. Nadeem, W. Lee, and C. Sodini, A higher-order topology for interpolative modulators for oversampling A/D converters. *IEEE Trans. Circuits Syst.*, CAS–37, 309–318, March 1990.
17. R. Carley, A noise-shaping coder topology for 15+ bit converters. *IEEE J. Solid-State Circuits*, SC–24, 267–273, April 1989.
18. R. Walden et al., Architectures for higher-order multibit sigma-delta modulators, *Proceedings of the IEEE Int. Symp. Circuits Syst.*, Portland, Oregon, pp. 895–898, 1990.
19. B. Leung and S. Sutarja, Multibit sigma-delta A/D converter incorporating a novel class of dynamic element matching techniques. *IEEE Trans. Circuits Syst.*, CAS-39(1), 35–51, 1992.

Further Information

Max Hauser provides an extremely good overview of oversampling converters in ''Principles of oversampling A/D conversion,'' *J. Audio Eng. Soc.*, 39(1/2), 3–26, 1991.

Sources for further reading about Nyquist rate converters include D.J. Dooley, *Data Conversion Integrated Circuits*. New York: IEEE Press, 1980, and R.J. van der Plassche, *Integrated Analog-to-Digital and Digital-to-Analog Converters*, Boston: Kluwer Academic, 1994.

IEEE Journal of Solid-State Circuits, particularly the December issues, and the *IEEE Transactions on Circuits and Systems* are good sources for more recent research on data converters.

Index

A

Amplifier impulse response, 6-15
Amplifier transfer function, 6-18
Analog circuit cells
  balanced differential amplifier
    engineering constraints, 2-45–2-47
    generalized system diagram, 2-44–2-45
    monolithic fabrication process, 2-45
    single-ended output voltage, 2-44
  common-base amplifier
    circuit broadbanding, 2-32
    common-emitter–common-base cascode, 2-30–2-32
    current buffering purpose, 2-26–2-27
    diode resistance, 2-28
    driving point input resistance, 2-29
    driving point output resistance, 2-29–2-30
    equivalent circuit, 2-27–2-28
    Kirchhoff's current law constraint, 2-27
    Miller multiplication, 2-32
    Norton transconductance, 2-31
    NPN and PNP AC schematic diagrams, 2-26–2-27
    small-signal analysis, 2-27
    voltage divider, 2-31
  common bipolar junction transistor (BJT) biasing circuits
    cascode current mirror, 2-6
    improved bandgap reference, 2-9–2-10
    low-bias current mirror, 2-5–2-6
    simple band-gap reference, 2-9
    simple current mirror, beta helper, 2-2–2-3
    simple current mirror, emitter degeneration, 2-4
    V_{BE} multiplier circuit, 2-8–2-9
    V_{BE} referenced current mirror, 2-6–2-8
    well-defined temperature coefficient, 2-9
    Widlar current mirror, 2-4–2-5
    Wilson current mirror, 2-3–2-4
    Zener diode reference, 2-8
  common-collector amplifier
    AC schematic diagrams, 2-33
    active load, 2-35–2-36
    driving point input and output resistances, test circuit, 2-34
    Thévenin load resistance, 2-37
    voltage gain, 2-33
  common-emitter amplifier
    active current source load, 2-24–2-26
    driving point output resistance, 2-23
    macromodel, 2-19–2-20, 2-23
    Norton current and equivalent circuit, 2-19
    NPN and PNP AC schematic diagrams, 2-17, 2-20–2-21
    operation, 2-17
    output coupling capacitance, 2-24
    small-signal test structure, 2-18–2-19
    Thévenin source voltage, 2-18
  common-mode input voltage, 2-41–2-42
  Darlington connection
    forward transconductance, 2-41
    schematic diagram, 2-38–2-39
    small-signal equivalent circuit, 2-39–2-40
    transconductance amplifier, 2-38
    transconductance frequency response, 2-38–2-39
    voltage gain, 2-40–2-41
    Wilson mirror load, 2-39
  differential input source voltage, 2-41

diode-connected transistor
KVL analysis, 2-15
small-signal transistor model, 2-14–2-15
static common-base current gain, 2-15–2-16
subject diagram, 2-14
V_{BE} multiplier, 2-15–2-17
volt–ampere characteristics, 2-14
performance index, 2-43
small-signal model, 2-11–2-14
system-level diagram, differential amplifier, 2-41–2-42
Thévenin equivalent I/O circuits
balanced bipolar differential amplifier, 2-52–2-53
driving point common-mode output resistance, 2-49–2-50
driving point differential-mode output resistance, 2-49–2-50
Kirchhoff's current and voltage laws, 2-47
open-circuit differential-mode gain, 2-51
pertinent test cell, 2-48–2-49
test circuit, 2-47–2-49
Thévenin model, 2-51
two-port model, 2-54
zero input signal excitation, 2-49
Analog integrated circuits
chip parasitic
circuit model, 1-143
effects on circuits, 1-136–1-137
inductance, 1-139
measurement, 1-145
modeling technique, 1-135
nonlinear interconnects, 1-141–1-142
overlap capacitance, 1-133–1-134
packaged IC, 1-142–1-143
parallel line capacitance, 1-133–1-134, 1-136
PSPICE ac simulation, 1-143–1-144
resistance, 1-137–1-139
substrate capacitance, 1-133–1-134
transmission line behavior
effects on circuits, 1-140–1-141
modeling, 1-140
types, 1-139–1-140
Analog-to-digital converters (ADCs)
Nyquist rate type
1-clock type, 10-33–10-40
N-clock type, 10-40–10-46
2^N-clock type, 10-46–10-47
oversampled type
feedback path delay integrator, 10-48
first-order noise-shaping converter, 10-48–10-49
forward path delay integrator, 10-48
higher resolution, 10-47
linear system model, 10-47–10-48
magnitude response, 10-49
noise-shaping type, 10-47–10-48
oversampling ratio (OSR), 10-47
pattern noise, 10-50
SNR, 10-51
test techniques
crossplot technique, 10-32–10-33
differential nonlinearity error, 10-32
gain error, 10-30
integral nonlinearity error, 10-31
offset error, 10-31
transfer characteristics, 10-29–10-30
Anceau's PLL scheme, 9-64–9-65
Avalanche noise, 3-52–3-53

B

Backward difference transformation, 5-21
Balanced binary tree (BBT), 9-59–9-61
Barrel shifter, 9-76–9-77
BiCMOS amplifiers, 2-95–2-96
Bilateral *z*-transform (BZT), 7-30–7-31
Bipolar integrated circuit design, 2-1
Bipolar junction transistor
current gain
base–emitter voltage, 1-4, 1-6
base transport and emitter injection efficiency, 1-6
gain–current relationship, 1-6–1-7
Ebers–Moll model, 1-2–1-3
Gummel–Poon model, 1-4–1-5
high-current phenomena, 1-7–1-8
integrated NPN transistor, 1-11–1-12
lateral and vertical PNP transistor, 1-12
second-order effects, 1-14–1-15
SiGe HBTs
collector and base currents *vs.* EB voltage, 1-17
cutoff frequency *vs.* collector current, 1-18
energy band diagram, 1-17
industry practice and fabrication technology, 1-19–1-20
measured doping and Ge profile, 1-17–1-18
operation principle and performance advantage, 1-18–1-19
small-signal model, 1-9–1-10
SPICE model, 1-15–1-16
thermal sensitivity, 1-13–1-14
Bipolar noise
avalanche noise, 3-52–3-53
burst noise–RTS noise, 3-51–3-52
$1/f$ noise, 3-50–3-51
generation–recombination noise, 3-49–3-50
noise characterization
equivalent noise resistance and noise temperature, 3-54–3-55
equivalent noise voltage and current, 3-53–3-54
noise figure, 3-55
noise $1/f^2$, 3-51
shot noise, 3-49
thermal noise, 3-48–3-49
Bipolar transistor, 3-53–3-54

Broadband bipolar networks
bipolar transistor modeling, high frequency
hybrid π model, 3-2–3-3
Miller approximation, 3-3
modified equivalent circuit, 3-4
simplified high-frequency model, 3-3
split current source, 3-4
broadband amplifier stability
classical feedback system review, 3-31
criteria, 3-31–3-32
grounded capacitor compensation, 3-35–3-36
high-frequency performance, 3-36
input capacitance to ground inversion, 3-40
Miller compensation and pole separation, 3-33
op-amp internal compensation strategy, 3-31
phase lag neutralization, 3-38–3-39
power supply impedance, 3-37
resistive and capacitive load effects, 3-37–3-38
single-stage op-amp compensation, 3-34–3-35
two-stage op-amp architecture, 3-32–3-33
C_μ neutralization, 3-9–3-10
current conveyor
applications, 3-12–3-13
broadband analog amplifier, 3-12
current–voltage transfer relationship, 3-12
first-generation current conveyor (CCI), 3-12, 3-15–3-16
second-generation current conveyor (CCII), 3-12, 3-16–3-18
single BJT CCII–, 3-13–3-14
supply-current sensing, voltage op-amp, 3-14–3-15
current-feedback operation amplifier
analysis, 3-22–3-23
architecture, 3-18–3-19
basic current mirror, 3-26–3-28
closed-loop inverting operation, 3-21–3-22
closed-loop noninverting operation, 3-20–3-21
design and development, 3-18
differential-mode operation, 3-19–3-20
feedback current, 3-20
high-speed performance, 3-18
improved broadband current mirror, 3-28–3-29
input stage, 3-44–3-45
phase linearity, 3-29
pole frequency comparison, 3-23–3-24
practical considerations, broadband design, 3-30–3-31
R_2 value, 3-29–3-30
slow rate, 3-24–3-25
wideband and high-gain, 3-25–3-26
Miller's theorem, 3-2
negative feedback, 3-10–3-11
RF bipolar transistor layout, 3-11
single-gain stages
common-base (CB) stage, 3-8–3-9
common-collector (CC) stage, 3-6–3-8
common-emitter (CE) stage, 3-4–3-6
transfer function and bandwidth characteristic
current-feedback, 3-41–3-43
voltage-feedback, 3-43–3-44
Widlar current mirror, 3-45–3-47
Bulk–drain depletion capacitance, 1-54
Bulk-induced modulation, 1-37
Bulk transconductance, 1-60
Butterfly operations, 7-14

C

Cascade voltage switch logic (CVSL) gate, 9-80–9-81
Cauer form, 6-13
CC–CE stage, 3-10
CE–CB cascode stage, 3-9
Channel length modulation effect
drain current equation, 1-88–1-89
JFET technology and device, 1-87
super MOS transistor, 2-90
Circular convolution theorem, 7-11–7-12
Circular spiral inductors
concentric type, 1-127
mutual inductance, 1-129–1-130
self-inductance, 1-127–1-129
total inductance, 1-127
Class *AB* current conveyor, 3-15
Classical feedback system, 3-51
Clocked CMOS logic (C^2MOS), 9-78–9-79
Clocking schemes
current-steered logic and dedicated third-layer, 9-69
distribution
Anceau's PLL scheme, 9-64–9-65
balanced binary tree (BBT), 9-59–9-61
clock ring configuration, 9-56–9-57
clock trunk concepts, 9-53–9-56
delay, skew, and rise time comparison, 9-59
four-quadrant approaches, 9-50
Grover's interval-halving PLL scheme, 9-66–9-68
H-trees, 9-57–9-59
phase-locked loops, 9-61–9-63
PLLs, CMOS, 9-63–9-64
single-driver configurations, 9-48–9-49
symmetric and generalized clock buffer trees, 9-50–9-53
tuning, large systems, 9-68–9-69
mixed technology, 9-69

optoelectronic clock distribution, 9-69–9-70
principles
clock signal manipulation, 9-46–9-47
controlled skew introduction, 9-45–9-46
delay minimization, 9-47–9-48
dynamic logic, 9-42–9-44
flip-flops, 9-42
isochronic/equipotential regions, 9-35–9-36
latches, 9-41–9-42
multiple-phase overlapping clocks, 9-40
overlapping clock phase generator, 9-40–9-41
single-phase type, 9-37–9-38
skew and delay, 9-34–9-35
skew on-chip nature, 9-36–9-37
synchronizers and metastability, 9-44–9-45
two-phase clock generator circuit, 9-39–9-40
two-phase type, 9-38–9-39
$\overline{Q}$ elimination, 9-69
reconfigurable clock nets, 9-69–9-70
voltage swing, 9-69
Clock recovery architecture
early–late block diagram, 5-28–5-29
edge-detection-based method, 5-29
integrator, 5-28–5-29
quadricorrelator, 5-29–5-30
waveforms, 5-28–5-29
Closed-loop transfer function, PLL, 5-3–5-4
CMOS
amplifier, 2-94–2-95
mixer topology
linear MOS mixers, 4-22–4-23
nonlinearity and LO-feedthrough analysis, 4-23–4-25
switching modulators, 4-21–4-22
RF integration, 4-11
Code-error calibration, DAC, 10-17
Common-base (CB) stage, 3-8–3-9
Common-collector (CC) amplifier
AC schematic diagram, 2-33
active load, 2-35–2-36
driving point input and output resistance, test circuit, 2-34
equivalent circuit, 3-6
equivalent high-frequency model, 3-8
output impedance, 3-7–3-8
Thévenin load resistance, 2-37
transfer function, 3-7
transform property, 3-8
voltage gain, 2-33
Common-emitter (CE) stage
equivalent circuit model, 3-5
high-frequency model, 3-4
Miller approximation, 3-4–3-6
right-hand-plane (RHP) zero, 3-6
second-order characteristic equation, 3-6
Complementary bipolar technology (CBT), 2-6
Complementary pass-transistor logic (CPL) gate, 9-80
Composite super NMOS transistors, 2-91–2-92
Compound semiconductor FET technologies
HEMT device
cross section structure, 1-99
drain current–drain voltage characteristic, GaN, 1-98–1-99
efficiency, 1-99
GaAs MESFET, 1-95–1-96
gate connected field plate, GaN, 1-100
microwave and mm-wave performance, 1-97
microwave power amplifier performance, 1-100
recessed gate AlGaAs/GaAs structure, 1-96
Schottky gate characteristics, 1-96
heterojunctions, AlGaAs and GaAs transition, 1-94–1-95
III–V compound semiconductors, 1-92–1-93
wide bandgap compound semiconductors, 1-97–1-98
Continuous-time Fourier transform
arbitrary aperiodic signal, 7-15
frequency-selective behavior, 7-17–7-18
properties, 7-16
spectral density magnitude, 7-16–7-17
spectral density, sinusoidal pulse, 7-17
Current-controlled current source (CCCS), 1-71–1-72
Current-controlled voltage source (CCVS), 1-72
Current conveyor
applications, 3-12–3-13
broadband analog amplifier, 3-12
current–voltage transfer relationship, 3-12
first-generation current conveyor (CCI), 3-12, 3-15–3-16
second-generation current conveyor (CCII), 3-12, 3-16–3-18
single BJT CCII–, 3-13–3-14
supply-current sensing, voltage op-amp, 3-14–3-15
Current crowding effect, 1-8
Current-feedback operation amplifier
analysis, 3-22–3-23
architecture, 3-18–3-19
basic current mirror, 3-26–3-28
closed-loop inverting operation, 3-21–3-22
closed-loop noninverting operation, 3-20–3-21
design and development, 3-18
differential-mode operation, 3-19–3-20
feedback current, 3-20
high-speed performance, 3-18
improved broadband current mirror, 3-28–3-29
phase linearity, 3-29
pole frequency comparison, 3-23–3-24
practical considerations, broadband design, 3-30–3-31
R_2 value, 3-29–3-30

slow rate, 3-24–3-25
wideband and high-gain, 3-25–3-26
Current-ratioed DAC, 10-8–10-11

D

Data converters
analog-to-digital converters (ADCs)
Nyquist rate type, 10-33–10-47
oversampled type, 10-47–10-57
test techniques, 10-29–10-33
digital-to-analog converters (DACs)
architectures, 10-6–10-12
design issues, 10-3–10-5
error sources, 10-22–10-25
high resolution techniques, 10-13–10-22
low-spurious design examples, 10-26–10-28
signal-to-noise ratio and dynamic range, 10-2
transfer characteristics, 10-1–10-2
Delay-locked loop (DLL)
block diagram, 5-30
modern digital system, synchronous communication, 5-31
PLL bandwidth control, 5-26
timing relationships, 5-31–5-32
voltage-controlled delay line (VCDL), 5-30
Differential amplifier
C_μ neutralization, gain stage, 3-10
voltage gain cells, 2-96–2-97
Differential nonlinearity (DNL), 10-3
Diffused resistors
avalanche breakdown mechanism, 1-107
isolation region, 1-106
n^+ diffusion layer, 1-108–1-109
normalized frequency response, 1-107–1-108
n-type emitter-diffused resistor, 1-107–1-109
p-type resistor and n-type epitaxial (epi) region, 1-106–1-107
Digital circuits and systems
architecture, 9-33
clocking schemes
current-steered logic and dedicated third-layer, 9-69
distribution, 9-48–9-69
mixed technology, 9-69
optoelectronic clock distribution, 9-69–9-70
principles, 9-33–9-48
$\overline{Q}$ elimination, 9-69
reconfigurable clock nets, 9-69–9-70
voltage swing, 9-69
microprocessor-based design
architecture, 9-92–9-95
features, 9-83–9-86
with general purpose microprocessor, 9-95–9-100
guidelines, 9-110–9-111
interfacing, 9-100–9-105
memory, 9-86–9-92
with microcontroller, 9-105–9-110
MOS logic circuits
CMOS inverter, 8-8–8-10
digital inverter, 8-4–8-6
dynamic CMOS logic gates, 8-14–8-15
MOSFET models, 8-1–8-4
nMOS logic gates, 8-6–8-8
static CMOS logic gates, 8-11–8-13
MOS storage circuits
dynamic charge storage, 9-72–9-75
dynamic CMOS logic, 9-78–9-82
shift register, 9-75–9-77
programmable logic devices (PLDs)
combinational logic (CL) PAL devices, 9-11–9-16
combinational logic (CL) PLD classification, 9-10
complexity device ladder, 9-1–9-2
design process, 9-24–9-27
FPGA architectures, 9-22–9-25
notation, 9-4–9-5
programmable array logic (PAL), 9-8–9-10
programmable logic array (PLA), 9-5–9-6
programmable macrocell outputs, 9-18–9-21
programmable read only memory (PROM), 9-7–9-8
sequential PAL devices, 9-16–9-18
state machines synthesis, 9-30–9-31
technologies, 9-2–9-4
VHDL synthesis style, FPGA, 9-27–9-30
systolic arrays
concurrency, parallelism, pipelining, 9-111–9-117
digital filters, 9-117–9-122
eigenvalue and SVDs, 9-137–9-142
Kalman filtering (KF), 9-132–9-137
recursive LSs estimation, 9-126–9-132
systolic word and bit-level designs, 9-122–9-126
transmission gates (TG)
analog processing, 8-24–8-33
complementary transistor version, 8-17–8-18
continuous time filters, 8-27–8-28
digital processing, 8-15–8-24
MOS operational amplifier compensation, 8-24–8-26
pass-transistor logic, 8-18–8-24
single transistor version, 8-15–8-17
switched-capacitor circuits, 8-28–8-33
transimpedance compensation, 8-26–8-27
Digital signal processing
bilateral z-transform (BZT), 7-30–7-31
continuous-time Fourier transform
arbitrary aperiodic signal, 7-15
frequency-selective behavior, 7-17–7-18

properties, 7-16
sinusoidal pulse, spectral density, 7-17
spectral density magnitude, 7-16–7-17
convolution operation, 7-26
discrete Fourier transform (DFT)
aliasing error, 7-9–7-11
applications, 7-11–7-12
discrete Fourier series (DFS) pair, 7-8
Hann window, 7-11
inverse DFT (IDFT), 7-7–7-8
leakage error, 7-10
magnitude spectrum, bandlimited signal, 7-8
periodic discrete-time signal, 7-6
quantization error, 7-6
rectangular window, 7-11
sampling frequency, 7-9
spectral properties, continuous-time periodic signal, 7-8
X(k) magnitude spectrum, 7-9–7-10
discrete-time Fourier transform (DTFT)
Dirac impulse function, 7-20
periodic function and FS coefficients, 7-19
properties, 7-23
sampling process model, 7-18, 7-20
discrete-time signals, 7-5
fast Fourier transform, 7-13–7-14
Fourier series, continuous-time periodic signals
approximation error, 7-2
complex FS coefficients, 7-3–7-4
cyclical phenomena, 7-4
Dirichlet conditions, 7-2
Euler's identity, 7-3
fundamental frequency, 7-2
Gibbs oscillation, 7-4–7-5
mean square error minimization, 7-2
periodic signal, 7-1–7-2
frequency response, 7-26–7-28
ideal digital filters, 7-29–7-30
linear and time invariant discrete-time systems
block diagram, 7-23–7-24
characteristic equation, 7-25
initial conditions, 7-24–7-25
off-line processing, 7-24
unit pulse response, 7-25–7-26
sampling theorem
aliasing error, 7-21–7-22
antialiasing filter, 7-21–7-22
data reconstruction formula, 7-21
DTFT magnitude, 7-20–7-22
exponentially weighted sinusoidal pulse, 7-22
FT magnitude, 7-22
low-pass filter magnitude frequency response, 7-21
sampling frequency, 7-20–7-21
spectral density magnitude, bandlimited signal, 7-20
stability, 7-26
transfer function, 7-33
unilateral *z*-transform, 7-33–7-35
z-plane, 7-31–7-32
Digital-to-analog converters (DACs)
architectures
current-ratioed type, 10-8–10-11
R+C/C+R combination type, 10-11–10-12
resistor-string type, 10-6–10-8
design issues
conversion speed, 10-5
linearity, 10-3
monotonicity, 10-3–10-5
error sources
glitch, 10-22–10-23
noise, 10-25
timing error-word clock jitter, 10-23–10-24
voltage reference, 10-24–10-25
high resolution techniques
dynamic matching type, 10-13–10-14
electronic calibration type, 10-14–10-18
interpolative oversampling type, 10-18–10-22
low-spurious design examples
dynamic linearity enhancement, 10-27–10-28
self-trimming, 10-26–10-27
spatial averaging, 10-26
signal-to-noise ratio and dynamic range, 10-2
transfer characteristics, 10-1–10-2
Digital truncation errors, 10-18
Diode equation, 1-2
Direct code mapping, 10-15
Direct frequency synthesizer, 4-17
Discrete-time index, 7-4
Discrete-time transfer functions, 5-21
Down converter, 4-15–4-16
Drain–source channel resistance, 1-60

E

Ebers–Moll model, 1-2–1-3
Edge detection, NRZ data, 5-29
Electronic calibration type, DAC
capacitor ratio error measurement cycles, 10-16
code-error calibration, 10-17
current difference measurement cycles, 10-16–10-17
digital truncation errors, 10-18
direct code mapping, 10-15
self-calibration for individual capacitor errors, 10-15–10-16
Elmore's approach, 6-15
Emitter follower, *see* Common-collector (CC) amplifier
Epitaxial resistors, 1-111–1-112
Extended RF transistor model, 4-9–4-10

F

Fast Fourier transform (FFT), 7-13–7-14
FET/current mirror bias circuit, 2-63
Fialkov condition, 6-12
Field programmable gate arrays (FPGAs)
architectures
high-level layout, 9-22
LUT, 9-24
minimal CLB, 9-23–9-24
programmable elements, 9-23

XC4010XL chip, CLB, 9-24–9-25
technologies, 9-22
VHDL synthesis style
combinational logic, 9-29
latches, 9-30
registers and flip-flops, 9-27–9-29
Finite impulse response (FIR) filters, 7-28, 9-117–9-118
First-order infinite impulse response filter, 5-22
1/*f* noise, 3-50–3-51
Folded cascode amplifiers, biasing circuits, 2-72–2-73
Forward transconductance, 1-60
Fourier series, 5-7
Fractional-*N* synthesizers, 4-19–4-20
Frequency locked-loop (FLL), 5-9
Frequency synthesizer
block diagrams, 5-32–5-33
dual-modulus prescalar, 5-33–5-34
output frequency, 5-32

G

Gain-boosting principle, 2-90–2-91
Generation–recombination noise, 3-49–3-50
Generation–recombination phenomena, 1-4
Gibbs oscillation, 7-4–7-5
Grover's interval-halving PLL scheme, 9-66–9-68
Gummel–Poon model, 1-4–1-5

H

Hestenes algorithm, 9-140–9-141
High-speed voltage buffer, 3-37

I

Ideal digital filters, 7-29–7-30
Indirect convolution method, 7-12
Infinite impulse response (IIR) filters, 7-28–7-29
bidirectional systolic arrays, 9-121
clustered look-ahead, 9-119–9-120
I–O relationship, 9-118
overlapped subfilter scheme, 9-122
scattered look-ahead, 9-120–9-121
systolic ring scheme, 9-122
Input-referred third-order intercept point (IIP3), 4-5
Integral nonlinearity (INL), 10-3
Integrated PNP transistors, 1-12
Integrator-type DACs, 10-4–10-5
Intermodulation distortion (IMD), 4-4–4-5
Interpolative oversampling technique
alternative 1 bit DAC sampling constant, 10-21–10-22
delta-sigma modulation, 10-18–10-19
dynamic range, 10-20
one-bit/multibit, 10-20
postfiltering requirement, 10-22
stability, 10-19–10-20
switched-capacitor 1 bit DAC/filter, 10-20–10-21
Ion-implanted resistors, 1-112–1-113

J

JFET technology and devices
channel-length modulation effect, 1-87
ion implanted silicon JFET, IC process, 1-91
large-signal model, drain current equations, 1-88–1-89
operating regions
cutoff and subthreshold current regions, 1-86–1-87
ohmic and pinch-off region, 1-85–1-86
static current–voltage characteristics, 1-85
small-signal model, 1-89–1-90
static *I–V* characteristics, 1-84
temperature effects, 1-87–1-88
JK-flipflop PD, 5-13–5-14
Junction capacitors
abrupt pn junction, 1-115–1-116
base–collector capacitor structure, 1-117–1-118
base–emitter capacitor structure, 1-117–1-119
depletion width, 1-116–1-117
permittivity, 1-116

K

Kalman filtering (KF)
Faddeev algorithm, 9-134–9-137
model, 9-132–9-133
other forms, 9-133–9-134
systolic matrix implementation, 9-134
Kirchhoff's current law, 1-3, 3-21
Kirchhoff's voltage law, 2-5
Kirk effect, 1-7–1-8

L

Linear bipolar technology, canonic cells
balanced differential amplifier
engineering constraints, 2-45–2-47
generalized system diagram, 2-44–2-45
monolithic fabrication process, 2-45
single-ended output voltage, 2-44
common-base amplifier
circuit broadbanding, 2-32
common-emitter–common-base cascode, 2-30–2-32
current buffering purpose, 2-26–2-27
diode resistance, 2-28
driving point input/output resistance, 2-29–2-30
equivalent circuit, 2-27–2-28
Kirchhoff's current law constraint, 2-27
Miller multiplication, 2-32
Norton transconductance, 2-31
NPN and PNP AC schematic diagrams, 2-26–2-27
small-signal analysis, 2-27
voltage divider, 2-31

common-collector amplifier
AC schematic diagram, 2-33
active load, 2-35–2-36
driving point input and output resistance, test circuit, 2-34
Thévenin load resistance, 2-37
voltage gain, 2-33
common-emitter amplifier
active current source load, 2-24–2-26
driving point output resistance, 2-23
macromodel, 2-19–2-20, 2-23
Norton current and equivalent circuit, 2-19
NPN and PNP AC schematic diagrams, 2-17, 2-20–2-21
operation, 2-17
output coupling capacitance, 2-24
small-signal test structure, 2-18–2-19
Thévenin source voltage, 2-18
common-mode input voltage, 2-41–2-42
Darlington connection
forward transconductance, 2-41
schematic diagram, 2-38–2-39
small-signal equivalent circuit, 2-39–2-40
transconductance amplifier, 2-38
transconductance frequency response, 2-38–2-39
voltage gain, 2-40–2-41
Wilson mirror load, 2-39
differential input source voltage, 2-41
diode-connected transistor
KVL analysis, 2-15
small-signal transistor model, 2-14–2-15
static common-base current gain, 2-15–2-16
subject diagram, 2-14
V_{BE} multiplier, 2-15–2-17
volt–ampere characteristics, 2-14
performance index, 2-43
small-signal model, 2-11–2-14
system-level diagram, differential amplifier, 2-41–2-42
Thévenin equivalent I/O circuits
balanced bipolar differential amplifier, 2-52–2-53
driving point common-mode/differential mode output resistance, 2-49–2-50
Kirchhoff's current and voltage laws, 2-47
open-circuit differential-mode gain, 2-51
pertinent test cell, 2-48–2-49
test circuit, 2-47–2-49
Thévenin model, 2-51
two-port model, 2-54
zero input signal excitation, 2-49
Linear convolution method, 7-12
Load capacitance neutralization, 3-38–3-39
Long-tail pair input transconductance, 3-24–3-25
Loop filter (LF) gain, 5-8
Low-noise amplifier (LNA)
cutoff frequency, 4-14
drain current equation, 4-14
IIP2 and IIP3 plots, 4-15–4-16
noise and source impedance matching, 4-14
noise figure, 4-12–4-14
third-order intermodulation, 4-14
topology, 4-12–4-13
transit voltage, 4-15
Low-pass filter (LPF)
PLL techniques, 5-2–5-3
two-stage op-amp architecture, 3-32

M

Matched transistors, 2-2–2-3
Metal–oxide–silicon field effect transistor (MOSFET) technology
charge storage, 1-24–1-25
cutoff regime, 1-38
depletion capacitance, 1-53–1-54
depletion zone analysis
body effect voltage, 1-35–1-36
charge density, 1-33–1-34
electric field intensity, 1-34–1-35
Gauss' law, 1-33
volt–ampere characteristics, 1-32–1-33
design-oriented analysis strategy
biasing, 1-77–1-78
circuit structure, 1-78
comments, 1-80–1-81
forward static transfer characteristic, 1-79–1-81
parameterization process, 1-79
transconductance coefficient, mobility degradation, 1-77–1-78
gate–bulk capacitance
characteristics, 1-32–1-33
depletion layer, 1-31
N-channel MOSFET, 1-30
pertinent equivalent circuit, 1-29
silicon dielectric constant, 1-29, 1-31
surface capacitance density, 1-29
Kirchhoff 's voltage law, 1-24
large-signal model, 1-56–1-57
lateral electric fields
carrier mobility degradation, 1-48–1-49
carrier velocity, 1-46–1-47
critical electric field, 1-46
drain saturation voltage, 1-48
Level 49 HSPICE model, 1-51
modulation voltage, 1-48
voltage and current correction factors, 1-49–1-50
ohmic regime
channel potential, 1-39–1-40
cross section, 1-38–1-39
drain saturation voltage, 1-38
pinched off channel, 1-40–1-41
static circuit model, 1-41–1-42
transconductance coefficient, 1-40

saturation regime
built-in potential, 1-44
channel length modulation voltage, 1-43
common-source volt–ampere characteristic curves, 1-45
drain current, 1-42, 1-44
large-signal circuit model, 1-44
strong inversion, 1-42–1-43
small-signal model
analysis, 1-58–1-59, 1-71
bulk-gate transconductance, 1-68
bulk modulation factor, 1-62, 1-69
bulk transconductance, 1-60, 1-62
common-source interconnection, 1-70
equivalent circuit, 1-70
forward transadmittance, 1-71
forward transconductance, 1-60–1-62
HSPICE model, 1-66
N-channel MOSFET operation, 1-59–1-60
P-channel MOSFET operation, 1-59–1-61
radial signal frequency, 1-68
scattering parameters, 1-66–1-67
short circuit admittance parameters, y_{ij}, 1-66–1-67
signal drain current, 1-69
simulating results, 1-74–1-76
three-port network, 1-66
VCCS synthesis, 1-71–1-72
surface charge density
Boltzmann's constant, 1-26
equilibrium condition, 1-27
Fermi potential, 1-26, 1-28
Gauss' law, 1-27
magnitude, 1-28
surface electron concentration, 1-27–1-28
temperature effects, 1-52–1-53
threshold condition, 1-36–1-37
unity gain frequency
radio frequency choke (RFC), 1-63
small-signal equivalent model, 1-63–1-64
voltage-controlled current source (VCCS), 1-64–1-65
vertical electric fields, 1-51–1-52
Metal–oxide–silicon (MOS) technology
capacitors, 1-119–1-120
current bias circuit, 2-68
M68HC11 microcontroller
circuit diagram, 9-106
memory map and sample ROM content, 9-107–9-108
operational modes, 9-106
programming model, 9-105–9-106
three button and four-digit LCD display device, 9-109–9-110
timing diagram, 9-106–9-107
Microprocessor-based design
architecture
bit level activity, ALU, 9-93
data paths, 9-93–9-94
other registers, 9-93–9-94
register-to-register transfer activity, accumulator, 9-92
binary cell (BC)
2 bits/word programmable ROM, 9-88–9-89
cascade and parallel construction, RAM module, 9-88
circuit, 9-86
EEPROM, 9-91
8K × 8 EPROM and package pin assignment, 9-89–9-90
map, sample ROM content, and assembly language source, 9-91–9-92
module, 9-87–9-88
multiword read/write memory circuit, 9-86–9-87
PROM, 9-90
unidirectional to bidirectional bus conversion, 9-86–9-87
features
block diagram, 9-85–9-86
conceptual diagram, 9-83–9-84
with general purpose microprocessor (Z80)
clock cycles, 9-99–9-100
control signals tasks, 9-97–9-98
pin assignment, 9-96–9-97
programming model, 9-95–9-96
schematic diagram, 9-99
timing diagrams, 9-98–9-99
guidelines, 9-110–9-111
interfacing
D/A converter, 9-102
daisy chain, 9-104
I/O ports with handshaking, 9-103
opto-isolated power control circuits, 9-101–9-102
switches, 9-102
timer-counter circuit, 9-104–9-105
wire OR'd circuit, 9-104
write controlled A/D, 9-102
with microcontroller (M68HC11)
circuit diagram, 9-106
memory map and sample ROM content, 9-107–9-108
operational modes, 9-106
programming model, 9-105–9-106
three button and four-digit LCD display device, 9-109–9-110
timing diagram, 9-106–9-107
Miller effect, 3-9
Monolithic device models
bipolar junction transistor
base–emitter voltage, 1-4, 1-6
base transport efficiency, 1-6
collector and base currents *vs.* EB voltage, 1-17
cutoff frequency *vs.* collector current, 1-17–1-18
Ebers–Moll model, 1-2–1-3
emitter injection efficiency, 1-6

energy band diagram, 1-17
gain–current relationship, 1-6–1-7
Gummel–Poon model, 1-4–1-5
high-current phenomena, 1-7–1-8
industry practice and fabrication technology, 1-19–1-20
integrated NPN transistor, 1-11–1-12
lateral and vertical PNP transistor, 1-12
measured doping and Ge profile, 1-17–1-18
operation principle and performance advantage, 1-18–1-19
second-order effects, 1-14–1-15
small-signal model, 1-9–1-10
SPICE model, 1-15–1-16
thermal sensitivity, 1-13–1-14
HEMT devices
cross section structure, 1-99
drain current–drain voltage characteristic, GaN, 1-98–1-99
GaAs MESFET, 1-95–1-96
gate connected field plate, GaN, 1-100
microwave and mm-wave performance, 1-97
microwave power amplifier performance, 1-100
recessed gate AlGaAs/GaAs structure, 1-96
Schottky gate characteristics, 1-96
heterojunctions, AlGaAs and GaAs transition, 1-94–1-95
III–V compound semiconductors, 1-92–1-93
JFET technology and devices
channel-length modulation effect, 1-87
cutoff and subthreshold current regions, 1-86–1-87
ion implanted silicon JFET, IC process, 1-91
large-signal model, drain current equations, 1-88–1-89
ohmic and pinch-off region, 1-85–1-86
silicon, operation, 1-82–1-83
small-signal model, 1-89–1-90
static current–voltage characteristics, 1-85
static *I–V* characteristics, 1-84
temperature effects, 1-87–1-88
metal–oxide–silicon field effect transistor (MOSFET) technology
biasing, 1-77–1-78
body effect voltage, 1-35–1-36
carrier mobility degradation, 1-48–1-49
carrier velocity, 1-46–1-47
charge density, 1-33–1-34
charge storage, 1-24–1-25
circuit structure, 1-78
critical electric field, 1-46
cutoff regime, 1-38
depletion capacitance, 1-53–1-54
depletion layer, 1-31
drain saturation voltage, 1-48
electric field intensity, 1-34–1-35
forward static transfer characteristic, 1-79–1-81
Gauss' law, 1-33
Kirchhoff 's voltage law, 1-24
large-signal model, 1-56–1-57
Level 49 HSPICE model, 1-51
modulation voltage, 1-48
N-channel MOSFET, 1-30
parameterization process, 1-79
pertinent equivalent circuit, 1-29
radio frequency choke (RFC), 1-63
silicon dielectric constant, 1-29, 1-31
small-signal equivalent model, 1-63–1-64
surface capacitance density, 1-29
temperature effects, 1-52–1-53
threshold condition, 1-36–1-37
transconductance coefficient, mobility degradation, 1-77–1-78
vertical electric fields, 1-51–1-52
voltage and current correction factors, 1-49–1-50
voltage-controlled current source (VCCS), 1-64–1-65
volt–ampere characteristics, 1-32–1-33
wide bandgap compound semiconductors, 1-97–1-98
MOSFET biasing circuits
CMOS technology
BJT, 2-57–2-58
circuit design, 2-55
diffusion resistors realization, 2-58–2-59
integrated circuit implementation, 2-56
N-well and P-well processes, 2-58
principal devices, 2-56–2-57
voltage reference, 2-56
device models and parameter variability
proportional to absolute temperature (PTAT), 2-62
temperature dependence, 2-59, 2-61
threshold voltage, 2-60–2-61
dynamic biasing, 2-75–2-76
low power supply voltage, 2-74–2-75
N- and P-doped polysilicon gate threshold, 2-70
simple amplifiers and other circuits
current mirrors, 2-73
folded cascode amplifiers, 2-72–2-73
single-stage amplifier, cascode loads, 2-71–2-72
two-stage amplifier, 2-70–2-71
voltage and current references
bandgap principle, bipolar technology, 2-66
BJT V_{be}-based references, 2-64–2-65

BJT V_T-based references, 2-65–2-66
curvature-compensated bandgap references, 2-67
discrete time bandgap references, 2-67–2-68
enhancement and depletion-mode threshold voltage difference, 2-69–2-70
lateral bipolar devices, 2-68–2-69
MOSFET threshold, 2-63
operational amplifier, 2-66
subthreshold region, 2-68
supply-voltage, 2-62–2-63
MOSFET technology, canonical cells
BiCMOS amplifiers, 2-95–2-96
CMOS amplifier, 2-94–2-95
composite transistors
BiCMOS, 2-84–2-85
bidirectional, 2-86–2-87
body effect, threshold voltage, 2-84
diode leakage current, 2-89
drain current, 2-87–2-88
equivalent circuit, 2-86–2-87
I_D curves, 2-86, 2-88
MOS-folded composite transistors, 2-85–2-86
physical cross section, 2-86–2-87
simulation program with integrated circuit emphasis (SPICE) model, 2-86, 2-90
various bulk connections, 2-86–2-87
V_{S1} curves, 2-87–2-88
differential amplifier, 2-96–2-97
folded-cascode operational amplifier, 2-98
matched device pairs
composite MOSFET (COMFET) circuits, 2-79
current mirrors, 2-79
differential pairs, 2-78–2-79
drain and differential current, 2-78
operation, saturation region, 2-78–2-79
transistor pairs operation, triode region, 2-79–2-81
voltage follower, 2-78–2-79
NMOS amplifier, 2-92–2-94
super MOS transistors
composite super NMOS transistors, 2-91–2-92
output impedance, 2-90–2-91
regulated current mirrors, 2-90
simulated I_D curves, 2-91, 2-93
unmatched device pairs
CMOS composite transistor, 2-82–2-83
CMOS inverter, 2-82, 2-84
drain current, 2-82
parallel and series composite NMOS transistors, 2-81–2-82
MOS logic circuits
CMOS inverter
circuit and switch model, 8-8–8-9
power dissipation, 8-10
VTC, 8-9–8-10
digital inverter
switching times, 8-5–8-6
symbol and electronic parameters, 8-4
voltage transfer characteristics (VTC), 8-5
dynamic CMOS logic gates
charge sharing problem, 8-14–8-15
three-input NAND gate and timing intervals, 8-14
MOSFET models
capacitances, 8-2–8-3
primary device voltages, 8-1–8-2
switching models, 8-3–8-4
symbols, 8-3
nMOS logic gates
AOI gates, 8-7–8-8
configurations, 8-6–8-7
NOR and NAND gates, 8-7–8-8
threshold voltage loss, 8-7
static CMOS logic gates
AOI gate and XOR circuit, 8-11–8-12
NAND and NOR gates, 8-11–8-12
pseudo-nMOS logic circuits, 8-12
TG-based 2:1 multiplexer, 8-12–8-13
transmission gate (TG), 8-12–8-13
XOR and XNOR gates, 8-13
MOS storage circuits
dynamic charge storage
charge sharing, 9-74–9-75
nMOS–nMOS, 9-73
nMOS–pMOS, 9-73–9-74
pMOS–pMOS, 9-73
source–drain connection storage nodes, 9-72–9-73
source–gate connection, 9-74
dynamic CMOS logic
cascade voltage switch logic (CVSL) gate, 9-80–9-81
clocked CMOS logic (C^2MOS), 9-78–9-79
complementary pass-transistor logic (CPL) gate, 9-80
domino CMOS logic, 9-79–9-80
NORA, 9-81–9-82
precharge–evaluate logic gate, 9-78
shift register
clocked barrel shifter, 9-76–9-77
parallel type, 9-76
simple type, 9-75–9-76
MOS transistor, 4-15

N

Natural frequency, PLL, 5-23–5-24
N-channel MOSFET (NMOS), 1-22
Network realization, 6-4–6-5
NMOS amplifier
common-source, enhancement load, 2-92–2-93
gain, 2-94
small-signal equivalent circuit, 2-94
Noise spectral density, 3-50–3-51
Noise temperature, 3-55
Noninverting amplifier, voltage-feedback op-amp, 3-23–3-24
Non-return-to-zero (NRZ)
data format, 5-27–5-28
phase and frequency detectors, 5-15

NORA CMOS dynamic logic, 9-81–9-82
Nyquist equation, 3-48
Nyquist rate converters
1-clock type
digital-error correction, 10-38–10-40
flash/parallel type, 10-33–10-35
pipelined type, 10-36–10-38
subranging converters, 10-33–10-34, 10-36
N-clock type
algorithmic type, 10-43–10-46
self-calibration successive approximation type, 10-42–10-43
successive approximation type, 10-40–10-42
2^N-clock type, 10-46–10-47

O

Ohmic regime
channel potential, 1-39–1-40
cross section, 1-38–1-39
drain saturation voltage, 1-38
pinched off channel, 1-40–1-41
static circuit model, 1-41–1-42
transconductance coefficient, 1-40
Operational amplifier (op-amp), rail-to-rail operation, 2-74
Optoelectronic clock distribution, 9-69–9-70
Output-referred third-order intercept point (OIP3), 4-5
Oversampled converters, ADC
feedback path delay integrator, 10-48
first-order noise-shaping converter, 10-48–10-49
forward path delay integrator, 10-48
higher resolution, 10-47
linear system model, 10-47–10-48
magnitude response, 10-49
noise-shaping type, 10-47–10-48
oversampling ratio (OSR), 10-47
pattern noise, 10-50
SNR, 10-51

P

Passive components
circular spiral inductors
concentric type, 1-127
mutual inductance, 1-129–1-130
self-inductance, 1-127–1-129
total inductance, 1-127
conductivity, 1-103–1-104
diffused resistors
avalanche breakdown mechanism, 1-107
isolation region, 1-106
n^+ diffusion layer, 1-108–1-109
normalized frequency response, 1-107–1-108
n-type emitter-diffused resistor, 1-107–1-109
p-type resistor and n-type epitaxial (epi) region, 1-106–1-107
electron and hole mobility *vs.* impurity concentration, 1-104–1-105
epitaxial resistors, 1-111–1-112
ion-implanted resistors, 1-112–1-113
junction capacitors
abrupt pn junction, 1-115–1-116
base–collector capacitor structure, 1-117–1-118
base–emitter capacitor structure, 1-117–1-119
depletion width, 1-116–1-117
permittivity, 1-116
MOS capacitors, 1-119–1-120
n-type nonuniformly doped resistor, 1-104
parallel-plate capacitor, 1-114
pinched resistors, 1-110–1-111
polysilicon capacitors, 1-120–1-121
rectangular spiral inductors
electrical model, 1-124–1-125
mutual inductance, 1-122–1-124
self-inductance, 1-122–1-123
self-resonant frequency, 1-125
total inductance, 1-122
transformer structure, 1-125–1-126
sheet resistance, 1-103–1-104, 1-106
thin-film resistors, 1-113–1-114
uniformly doped resistor, 1-103
Passive lag filter, 5-4
P-channel MOSFET (PMOS), 1-22–1-23
Phase detector (PD), 5-1–5-2
Phase-error transfer function, 5-4
Phase-frequency detector (PFD), 5-2
Phase-locked loop (PLL) circuits
applications
clock recovery architecture, 5-28–5-30
data conversion, 5-28
data format, 5-27–5-28
delay-locked loop, 5-30–5-32
frequency synthesizer, 5-32–5-34
basic operation concepts, 5-1–5-2
charge-pump
block diagram, 5-22
closed-loop transfer function, 5-23
LF schematic, 5-22–5-23
normalized natural frequency, 5-23
purpose, 5-22
static phase error, 5-23–5-24
classification, 5-2
continuous-time loop filter (LF)
Bode plots, 5-16–5-17
frequency response, 5-17–5-18
high and low gain loop, 5-17
second-order loop *vs.* damping factors, 5-17–5-18
transfer function, 5-16–5-17
types, 5-16
definition, 5-1
design considerations
adaptive-bandwidth, 5-25

linear model, charge-pump PLL, 5-26
process, voltage, and temperature variations, 5-24–5-25
timing recovery, 5-26
typical procedures, 5-24
wide-range LC-tank VCO, 5-25
discrete-time loop filter (LF), 5-21–5-22
frequency synthesizer, 4-17–4-18
phase and frequency detectors
JK-flipflop, phase tracking, 5-13–5-14
operation, 5-15
waveforms, 5-15
XOR, 5-13–5-14
s-domain to *z*-domain transformations
backward difference method, 5-19–5-20
bilinear transformation method, 5-20–5-21
rectangular area approximation, 5-19–5-20
trapezoidal integration method, 5-20–5-21
techniques
acquisition process, 5-8–5-9
basic topology, 5-2–5-3
lock-in process, 5-7–5-8
loop orders, 5-3–5-4
noise performance, 5-9–5-10
tracking process, 5-4–5-6
voltage-controlled oscillators
harmonic oscillator, 5-11
resistive tuning, 5-12
ring oscillator topology, 5-11
Poisson's equation
JFET static *I–V* characteristics, 1-84
surface charge density, 1-26
Pole frequency
current-feedback operations amplifier, 3-23–3-24
two-stage op-amp architecture, 3-33
Polysilicon capacitors, 1-120–1-121
Potential divider, 2-62
Power added efficiency (PAE), 4-27
Power amplifier (PA)
CMOS RF amplifier
linearization techniques, 4-27–4-29
technology, 4-25
switching class E amplifier, 4-25–4-27
Prescaler, 4-19
Programmable array logic (PAL)
architecture, 9-8–9-9
example, 9-9–9-10
Programmable logic array (PLA)
architecture, 9-5
programmed chip, 9-6
Programmable logic devices (PLDs)
combinational logic (CL) classification, 9-10
combinational logic PAL devices
examples, 9-14–9-16
function implementation, 9-11, 9-13
implementation range, 9-14–9-15
PAL16L8 chip logic diagram, 9-11–9-12
complexity device ladder, 9-1–9-2
design process
design entry modes, 9-25–9-26
flow diagram, 9-24–9-26
logic synthesis, 9-26–9-27
mapping and simulation, 9-27
FPGA architectures
high-level layout, 9-22
LUT, 9-24
minimal CLB, 9-23–9-24
programmable elements, 9-23
XC4010XL chip, CLB, 9-24–9-25
FPGA technologies, 9-22
notation, 9-4–9-5
programmable array logic (PAL)
architecture, 9-8–9-9
example, 9-9–9-10
programmable logic array (PLA)
architecture, 9-5
programmed chip, 9-6
programmable macrocell outputs, PAL
circuit diagram, 9-18–9-19
circuit sizes, 9-21
macrocell architecture, 9-19–9-20
PAL22V10 chip architecture, 9-18
switch settings, 9-20–9-21
programmable read only memory (PROM)
conceptual diagram, 9-7
truth table, 9-7–9-8
sequential PAL devices
implementation range, 9-16, 9-18
PAL16R4 chip logic diagram, 9-16–9-17
state machines synthesis, 9-30–9-31
technologies
process type, 9-2–9-3
programming type, 9-3–9-4
VHDL synthesis style, FPGA
combinational logic, 9-29
latches, 9-30
registers and flip-flops, 9-27–9-29
Pull-in process, 5-8–5-9
Pulse-forming network output response, 6-7

Q

$\overline{Q}$ elimination, 9-69
Q-factor, 4-11–4-12

R

Radio frequency (RF) front-ends, 4-1–4-2
Radix-2 FFT algorithm, 7-13
R+C/C+R combination DAC, 10-11–10-12
Reactance pulse-forming network synthesis
delayed output pulse
algebraic ratio, 6-11
Cauer form, 6-13
delayed quasi-rectangular pulse, 6-9, 6-12
deterioration, 6-13–6-15
Fialkov condition, 6-12
Laplace transform, 6-10–6-11
normalized time, 6-9–6-10
pulse parameter, 6-12
slew rate, fronts, 6-9, 6-11

step response, 6-12–6-13
transfer function realization, 6-13
voltage source efficiency coefficient, 6-9–6-10, 6-13–6-14
non-delayed output pulses, 6-7–6-8
quasi-rectangular output pulse and laplace transform approximation, 6-4
impulse excitation, 6-2
impulse response, 6-4
rise and fall times, 6-3
shift theorem, 6-4
realization requirements, 6-4–6-5
second approximation step impulse response $V_o(s)$, 6-5
spectrum bandwidth ω_m, 6-6
transfer function, 6-5–6-6
sinusoidal pulse forming approximation, realization, 6-21–6-22
transfer function, 6-20–6-21
wideband amplifiers, transfer functions
approximation, 6-16–6-17
design, 6-18–6-19
Elmore's approach, 6-15
step response parameters and Laplace transform, 6-15–6-16
tabulated results, 6-19
Real feedback amplifier, 3-36
Realizable transfer function, 6-18
Rectangular spiral inductors
electrical model, 1-124–1-125
mutual inductance, 1-122–1-124
self-inductance, 1-122–1-123
self-resonant frequency, 1-125
total inductance, 1-122
transformer structure, 1-125–1-126
Region of convergence (ROC), 7-32
Relaxation oscillators, 5-11
Resistor/current mirror bias circuits, 2-63
Resistor-string DAC, 10-6–10-8
RF communication circuits
frequency synthesizer
fractional-*N* synthesis, 4-19–4-20
oscillator, 4-18–4-19
prescaler, 4-19
topology, 4-17–4-18
receiver
down converter, 4-15–4-16
LNA, 4-12–4-15
signal interference, 4-3
system performance metrics
first- and third-order intermodulation, 4-5
gain compression, 4-5
LNA, 4-4
noise figure, 4-3–4-4
nonlinearity performance, 4-4–4-5
receiver sensitivity, 4-3
signal-to-noise and distortion ratio (SNDR), 4-4
signal-to-noise ratio (SNR), baseband, 4-3–4-4
technology
active devices, 4-9–4-10
passive devices, 4-11–4-12
transceiver architectures
direct/zero-IF receiver, 4-7–4-8
down-conversion process, 4-6–4-7
heterodyne receiver, 4-6–4-7
image/mirror signal, 4-6
image rejection, 4-7–4-8
image suppression, 4-7
low output impedance driver, 4-6–4-7
quadrature amplitude modulation (QAM), 4-8
transmission medium, 4-2–4-3
transmitter
CMOS power amplification, 4-25
CMOS RF PA linearization, 4-27–4-29
linear MOS mixers, 4-22–4-23
nonlinearity and LO-feedthrough analysis, 4-23–4-25
switching class E amplifier, 4-25–4-27
switching modulators, 4-21–4-22
up *vs.* down conversion, 4-20–4-21
Ring oscillator jitter, 5-9–5-10

S

Saturation current, 1-4–1-5
Saturation regime
built-in potential, 1-44
channel length modulation voltage, 1-43
common-source volt–ampere characteristic curves, 1-45
drain current, 1-42, 1-44
large-signal circuit model, 1-44
strong inversion, 1-42–1-43
Second-generation current conveyor (CCII), 3-16–3-18
Segmented DACs, 10-4
Self-biased V_{BE} referenced current source, 2-7
Self-biased V_T referenced current source, 2-7–2-8
Self-calibration for individual capacitor errors, DAC, 10-15–10-17
Series R–C snubber, 3-39
Shot noise, 3-49
Signal interference, 4-3, 4-5
Signal-to-noise and distortion ratio (SNDR), 4-4
Simple current mirror, beta helper, 2-2–2-3
Simulation program with integrated circuit emphasis (SPICE) model
bipolar transistor parameters, 1-15–1-16
Gummel–Poon model, 1-5
small-signal model, 1-9
thermal sensitivity, 1-13–1-14
Single-stage op-amp architecture, 3-34–3-35
Sinusoidal envelope, output response, 6-20
Skew and delay, 9-34–9-35
Small-signal model
BJT
Boltzmann voltage, 2-12
emitter–base junction diffusion resistance, 2-13
equivalent circuit, 2-10–2-11
forward Early resistance and transconductance, 2-13
large-signal model, 2-11–2-12
monolithic fabrication process, 2-12

quiescent-operating point, 2-13
static common-emitter current, 2-12
MOSFET technology
analysis, 1-58–1-59, 1-71
bulk-gate transconductance, 1-68
bulk modulation factor, 1-62, 1-69
bulk transconductance, 1-60, 1-62
common-source interconnection, 1-70
equivalent circuit, 1-70
forward transadmittance, 1-71
forward transconductance, 1-60–1-62
HSPICE model, 1-66
N-channel MOSFET operation, 1-59–1-60
P-channel MOSFET operation, 1-59–1-61
radial signal frequency, 1-68
scattering parameters, 1-66–1-67
short circuit admittance parameters, y_{ij}, 1-66–1-67
signal drain current, 1-69
simulating results, 1-74–1-76
three-port network, 1-66
VCCS synthesis, 1-71–1-72
Square-law current model, 1-88, 1-90
Statz model, 1-89
Step response parameters, 6-19
Stray capacitance, 3-40
Stripe geometry, 3-11
Supply decoupling circuitry, 3-37
Synchronizers and metastability, 9-44–9-45
Systolic arrays
concurrency, parallelism, pipelining
definitions, 9-111–9-112
design techniques, 9-114–9-117
linear and rectangular type, 9-112
linear correlation, 9-113–9-114
uniprocessor system, 9-112
digital filters
finite impulse response (FIR) filters, 9-117–9-118
infinite impulse response (IIR) filters, 9-118–9-122
eigenvalue and SVDs
rectangular matrix, Hestenes algorithm, 9-140–9-141
rectangular nonsymmetric matrix, 9-141–9-142
spatial filtering problem, 9-137–9-138
symmetric matrix, 9-139
Kalman filtering (KF)
Faddeev algorithm, 9-134–9-137
model, 9-132–9-133
other forms, 9-133–9-134
systolic matrix implementation, 9-134
recursive LS estimation
Givens orthogonal transformation, 9-128–9-129
optimal residual and solutions, 9-129
QRD, 9-127
sliding window and forgetting factor approach, 9-127
technique, 9-126–9-127
triangular implementation, 9-129–9-132
systolic word and bit-level designs
advantages, 9-123
hierarchical approach, 9-122–9-123
serial convolver, 9-123–9-126

T

Table look-up synthesizer
fast Fourier transform, 7-13
topology, 4-17
Thermal noise, 3-48–3-49
Thermal voltage, V_T, 2-4–2-5
Thin-film resistors, 1-113–1-114
Tracking process, PLL
acceleration error, 5-5–5-6
final value theorem, 5-5
hold range, 5-6
Laplace transform, 5-5
step phase error, 5-5–5-6
velocity error/static phase error, 5-5
Transfer function poles, 6-19
Transmission gates (TG)
analog processing, 8-24–8-33
complementary transistor version
resistances, 8-18
voltage transmission properties, 8-17–8-18
continuous time filters, 8-27–8-28
digital processing, 8-15–8-24
MOS operational amplifier compensation
fully differential folded-cascode op-amp, 8-25–8-26
two-stage op-amp, 8-24–8-25
pass-transistor logic
adder, 8-21–8-22
CMOS *D* latch, 8-21
CPL circuit and modules, 8-22–8-23
CPL full adder circuit, 8-23
Karnaugh map, XOR function, 8-18–8-19
model, 8-18–8-19
OR gates, 8-20
16 ×16 bit multiplier, 8-23–8-24
SRAM and DRAM cells, 8-21
truth table, XOR function, 8-18–8-19
two-input multiplexer, 8-20
XOR gates, 8-19–8-20
single transistor version
I–O characteristics, 8-16–8-17
nMOS and pMOS, 8-15–8-16
switched-capacitor circuits
bottom-plate differential-input lossless digital integrator (LDI), 8-29–8-30
differential bilinear integrator, 8-29–8-30
direct digital integrator (DDI), 8-28–8-29
switch charge injection analysis, 8-30–8-33

transimpedance compensation
neuron and synapse operation, 8-27
preamplifier circuit, 8-26
Typical switched capacitor integrator, 2-75

U

Ultra-wideband (UWB) technology, 6-19–6-20
Unit step function, 7-5

V

VHDL, programming language
combinational logic, 9-29
latches, 9-30
registers and flip-flops
asynchronous reset code, 9-28
synchronous reset code, 9-27
synthesis tools, 9-28
state machine, 9-30–9-31
Voltage-controlled current source (VCCS), 1-64–1-65
Voltage-controlled oscillator (VCO)
frequency synthesizer, 4-17
LC-tank, 4-18
passive device, 4-12
RF transceiver architectures, 4-9
Voltage-controlled voltage source (VCVS), 1-71–1-72
Voltage follower, 3-14–3-15
V_t-referenced current bias circuit, 2-64

W

Wideband amplifiers, transfer functions
approximation, 6-16–6-17
design, 6-18–6-19
Elmore's approach, 6-15
step response parameters and Laplace transform, 6-15–6-16
tabulated results, 6-19
Wide bandgap compound semiconductors, 1-97–1-98
Widlar current mirror
bipolar junction transistor (BJT) biasing circuit, 2-4–2-5
current-feedback operations amplifier, 3-26–3-27
Wilson current mirror, 2-3–2-4

X

XOR PD, 5-13–5-14

Z

Z80 microprocessor-based design
clock cycles, 9-99–9-100
control signals tasks, 9-97–9-98
pin assignment, 9-96–9-97
programming model, 9-95–9-96
schematic diagram, 9-99
timing diagrams, 9-98–9-99